

661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
 b 2197 GCGAGGCAAGACTGCTGGCCGGCTCATGAGCGTGCAGATCCGCCACTTCTGCT 2256
 y 681: ProLeuLysGlnTrpPheAlaCyVal 689
 b 2257 CCGCTGAAGAGATTTTGGCTGTGTG 2283

RESULT 2

S-10-144-577-2

Sequence 2, Application US/10144577
 Publication No. US20030083292A1

GENERAL INFORMATION:

APPLICANT: Macleod, Alan Robert

TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms

FILE REFERENCE: MET-005

CURRENT APPLICATION NUMBER: US/10144,577

PRIOR FILING DATE: 2002-05-13

PRIOR APPLICATION NUMBER: US 60/290,202

PRIOR FILING DATE: 2001-05-14

PRIOR APPLICATION NUMBER: US 60/290,212

NUMBER OF SEQ ID NOS: 49

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 3005

TYPE: DNA

ORGANISM: Homo sapiens

S-10-144-577-2

Alignment Scores:

red. No.: 0

score: 3763.00

Percent Similarity: 100.0%

Best Local Similarity: 100.0%

Very Match: 100.0%

3: 6

Gaps: 0

Length: 3005

Matches: 689

Conservative: 0

Mismatch: 0

Indels: 0

Gaps: 0

S-10-623-813-86 (1-689) x US-10-144-577-2 (1-3005)

1 MetAenAlaValIGluGluGlnGlyProGlyGluSerGlnLysValGluGlnAlaSer 20
 899 ATGATGCTGTGGAAGAAACCAAGGAGCCCGGAGACTCTCAAGAGTGGAGAGGCGAC 958
 21 ProProAlaValIGlnGlnProThraProAlaSerProThraValAlaThraProGlu 40
 959 CCTCTGCTGTGCAAGCCCACTGATCCCGCATCCCACTGTGCTACACGCTGTAG 1018
 41 ProValIGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
 1019 CCGGTGGGTCCGATGCTGGGCAAGATGCCCAAGACGAGCATGACGAGCCAGAG 1078
 61 TyrGluAspGlyArgGlyPheGlyLysLysLysValThrGlyLysLeuArgGlyPhe 80
 1079 TACGAGGAGCGCGGGCTTGGCATTTGGGAGCTGTGTGGGGAATCTGGGGGCTTC 1138
 81 SerTrpTrpProGlyArgGlyLeuValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 1139 TCTCTGTGCGCCAGCGGCAATGTGTCTTGTGTGATACGGGCGGAGCCGAGCACTTGA 1198
 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCyValGluLys 120
 1199 GGCACCCGCTGGGTCTATGTGTTCCGAGACGGCAATTTCTCATGTGTGTGTGAGAG 1258
 121 LeuMetProLeuSerSerPheCySerAlaPheHisGlnAlaThrLysAlaGlnPro 140
 1259 CGATGCGCGTGAAGCTCTTTTCAAGTGGCTTCCACAGGCGCACTGACAAAGAGCC 1318
 141 MetTrpArgLysAlaIleTrpGlyValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 1319 ATGTACCGCAAGCCATCTTACGAGTCTTCCAGTGTGCGACGAGCCGCGGGGAGGCTG 1378

Qy 161 PheProValCyHisAspSerAspGluSerAspThrAlaLysAlaValGluGlnAsn 180
 Db 1379 TTCGGGTGTGCCACACAGCAAGATGACATCGCCAGGCGCTGTGAGTGCAGAAC 1438
 Qy 181 LysProMetIleGlnTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 Db 1439 AACCCATGATTAATGGCGCTTGGGGGCTTCCAGCTTCTGACCTTAAGGGCTTGGAG 1498
 Qy 201 ProProGluGluGluLysAsnProLysLysValLysLysValLysLysValLysLysVal 220
 Db 1499 CCACAGAGAGAGAGAGATCCCTCAAGAGAGTGTACCGGACATGTGGTGGAACT 1558
 Qy 221 GluAlaAlaAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 240
 Db 1559 GAGGAGCTGCTCCAC 1618
 Qy 241 LysProLysValLysGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 260
 Db 1619 AACCCAGCTCAAGAGATTTATGTAGTGGCGCACAGAGAGCGCTGTGTACAGGTG 1678
 Qy 261 ArgGlnLysCyAspArgAsnIleGluAspIleCyHisSerCyGlySerLeuAsnValThr 280
 Db 1679 CCGCAGAGATGCGGGAACATTTAGCATCTGATCTCTGTGGAGGCTCAATGTATAC 1738
 Qy 281 LeuGlnHisProLeuPheValGlyLysMetCyGlnAsnCyLysAsnCyPheLeuGlu 300
 Db 1739 CTGGAACACCCCTCTCTGTGAGGATGTGCCAAGCTGCAAGAACTGCTTCTGAGAG 1798
 Qy 301 CyAlaLysGlnLysAspAspAspGlyLysLysLysLysLysLysLysLysLysLys 320
 Db 1799 TGTGCTTACCAATGACGACACAGGCTACAGTCTTACTGACCACTGTGTGTGGGGG 1858
 Qy 321 ArgGluValLeuMetCyGlyLysAsnAsnAsnCyCyAspCyPheCyValGluCyVal 340
 Db 1859 CGTGAAGTGTCTATGTGGGGAACAACAACCTCTGAGGTGCTTTTGGTGAAGTGTGTG 1918
 Qy 341 AspLeuLeuValGlyProGlyValAlaGlnAlaAlaIleLysGluAspProTrpAsnCy 360
 Db 1919 GACCTTGTGTGGGCGCGGGGCTGCCGAGGACATTAAGAAAGACCTTGAACCTCG 1978
 Qy 361 TyrMetCyGlyHisLysGlyLysLysLysLysLysLysLysLysLysLysLysLys 380
 Db 1979 TACATGTGGGCAACAGGATCTTACGGGCTGTGCGGCGGAGAGGACTGGCCCTTC 2038
 Qy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValLys 400
 Db 2039 CCGCTCCAGATTTTCTGTGCTTAATACACAGCAAGATTTGACCTTCAAGGCTTAC 2098
 Qy 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyLys 420
 Db 2099 CCACCTGTCCAGCTGAG 2158
 Qy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyLysLysLysLysLysLysLysLys 440
 Db 2159 GCTACAGGCTCTGTGTCTGTAAGGACTTGGGCACTTCAAGTGGACCGCTTACCTTC 2218
 Qy 441 GluValCyGluAspSerLysLysLysLysLysLysLysLysLysLysLysLysLys 460
 Db 2219 GAGGTGTGAGGATCTCATTCAGGTGGCATGTGCGGACACAGGGAGAGATCATGTAC 2278
 Qy 461 ValGlyAspValArgSerValThrGlnLysHisLysLysLysLysLysLysLysLys 480
 Db 2279 GTGGGAGAGTCCGAGGCTTCAACAGAGATTCACAGAGTGGGCGCATTCGATCTG 2338
 Qy 481 ValIleGlyLysSerProCyAsnAspLeuSerLysLysLysLysLysLysLysLys 500
 Db 2339 GTGATTTGGGGGAGTCTCTGATGATGACCTTCACTGCAACCTGCTGCAAGGGCTC 2398
 Qy 501 TyrGluGlyThrGlyArgLeuPhePheGluPheLysLysLysLysLysLysLysLys 520
 Db 2399 TACGAGGAGTGGCGGCTCTTCTTGTAGTCTTACCGCTCTGATGATGATGATGATG 2458
 Qy 521 LysGluGlyLysAspAspArgProPhePheTrpLeuPheGluLysAsnValAlaMetGlyVal 540

Fri Oct 6 13:25:18 2006

us-10-623-813-8

Db 2459 AAGGAGGGAGATGATCGCCCTTCTTCTGGCTCTTTGAGAAATGTGGTGGCCATGGGCGTT 2518
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 2519 AGTGACAAGAGGGACATCTCGCGATTCTCGAGTCCAACCTGTGATGATGATGCCAAA 2578
QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db 2579 GAAGTGTGACCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAGG 2638
QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGluCysLeuGluHisGlyArg 600
Db 2639 CCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGGAGCATGGCAGG 2698
QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db 2699 ATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAGGTCAAACCTCCATAAAGCAGGGC 2758
QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db 2759 AAAGACCAGCATTTTCTGTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACCTGAA 2818
QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2819 ATGGAAAGGGTATTTGGTTTCCAGTCCACTATCTGACGTCTCCAACATGAGCCGCTTG 2878
QY 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2879 GCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCGCT 2938
QY 681 ProLeuLysGluTyrPheAlaCysVal 689
Db 2939 CCGCTGAAGGAGTATTTGCGTGTGTG 2965

RESULT 3

US-10-172-118-132

Commence 132, Application US/10172118

3022

| | | | |
|----|------|---|------|
| Db | 2039 | CGGGCTCCAGATGTTCTTCGCTAAATAACACAGACCAGAAATTGACCTCCCAAGGTTTAC | 2098 |
| Qy | 401 | ProProValProAlaGluValSarGlyProIleArgValIleuSerIleuPheAspGlyIle | 420 |
| Db | 2099 | CCACCTTCTCCAGCTGAGAAAGAGAGGCCATCCGGGTGCTGTCTCTTTATGAAATC | 2158 |
| Qy | 421 | AlaThrGlyIleuIleuValIleuLysAspIleuGlyIleGlnValAspArgTyrIleIleAser | 440 |
| Db | 2159 | GCTACAGAGGCTCCGTGGTGTGAAGAGACTTGGGCAATTCAGGTGACCGGTACATGGCTTCG | 2218 |
| Qy | 441 | GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr | 460 |
| Db | 2219 | GAGGTGTGTGGAGCTCCATCACGGTGGGATGTGGTGGGACCAAGGGGAAGATCATGTAC | 2278 |
| Qy | 461 | ValGlyAspValAlaXSerValThrGlnLysHisIleGlnGluIuTrpGlyProPheAspIleu | 480 |
| Db | 2279 | GTGGGGAGCTCCCGACGCTCACACAGAGATATCCAGAGTGGGGCCATTCGATCTCG | 2338 |
| Qy | 481 | ValIleGlyIleSerProCysAsnAspLeuSerIleValIleuProAlaArgLysGlyLeu | 500 |
| Db | 2339 | GTGATTTGGGGCAGCTCCGTGCAGATGACTCTCCATCTCAACCTGTCTCCAAAGGGCTTC | 2398 |
| Qy | 501 | TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuIleuHisAspAlaArgPro | 520 |
| Db | 2399 | TACGAGGGCACCTGGCGGCTCTCTTTAGATTCTTACGGCTCTCGCATGATGGCGGGCCC | 2458 |
| Qy | 521 | LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal | 540 |
| Db | 2459 | AAAGAGGAGATGATCGCCCTCTTCTGGCTCTTGAGATGTGGTGCGGCATGGGGGTT | 2518 |
| Qy | 541 | SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys | 560 |
| Db | 2519 | AGTGCACAAAGGAGACATCTCGGATTTCTCGAGTCCAACTGTGATGATGATGCCAA | 2578 |
| Qy | 561 | GluValSerIleAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg | 580 |
| Db | 2579 | GAAGTCTACGCTGCACACAGGGCCCGCTACTCTGTGGGGTAACCTTCCGGATGAAACAGG | 2638 |
| Qy | 581 | ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGluCysGluGlnHisGlyArg | 600 |
| Db | 2639 | CCGTTGGCATTCCACTGTGATGATTAAGCTGGAGCTGCAGAGATGTCTTGAACTAGCGAGG | 2698 |
| Qy | 601 | IleAlaLysPheSerLysValArgTrpIleThrTrpArgSerAsnSerIleGlyGlnGly | 620 |
| Db | 2699 | ATAGCCAAAGTTCAGCAAAAGAGAGACATTACTACAGAGGTCAAATCCATAAGCAGGGCC | 2758 |
| Qy | 621 | LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu | 640 |
| Db | 2759 | AAAGACCCAGCATTTTCTGTCTTCAATGATGAGAAAGAGACATCTTATGGTGCACTGAA | 2818 |
| Qy | 641 | MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu | 660 |
| Db | 2819 | ATGGAAGGGATATTGGTTTCCAGATCCACTATACGATGAGGTCTCCAAACATGAGCCGCTTG | 2878 |
| Qy | 661 | AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla | 680 |
| Db | 2879 | GCGAGGACAGACCTGCTGGGCGCGGTCACTGAGCGGTGCATCATCTCCGCCACTCTTCGCT | 2938 |
| Qy | 681 | ProLeuLysGluTyrPheAlaCysVal | 689 |
| Db | 2939 | CCGCTGAAGAGATATTTTGGCTGTGTG | 2965 |

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 3005)
Xie,S., Mang,Z., Okano,M., Nogami,M., Li,Y., He,W.W., Okumura,K.
and Li,E.
Cloning, expression and chromosome locations of the human DNMT3
gene family
Gene 236 (1), 87-95 (1999)
2 (bases 1 to 3005)
Xie,S., Okano,M. and Li,E.
Direct Submisison
Submitted (25-MAY-1998) CVMC, Mass. General Hospital, 149 13th
Street, Charlestown, MA 02129, USA
3 (bases 1 to 3005)
Hata,K., Shirohzu,H., Sasaki,H. and En,I.
Direct Submission
Submitted (12-FEB-2001) CVMC, Mass. General Hospital, 149 13th
Street, Charlestown, MA 02129, USA
Sequence update by submitter
On Feb 12, 2001 this sequence version replaced gi:4927369.

REMARK COMMENT FEATURES

source

1..3005 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p23; between D2S171 and D2S174; 48-50 CM"
1..3005 /gene="DNMT3A"
230..2968 /gene="DNMT3A"
/function="de novo DNA cytosine methyltransferase"
/codon_start=1
/product="DNA cytosine methyltransferase 3 alpha"
/protein_id="AAD33084.2"
/db_xref="GI:12746532"
translation="MPMAPSSGCGDITSSAAAREEDRKDGEEQEGRKEEROEEST
ARKRGRRGRKRKHPPEVSGDPDPDAVISKSPMAODSASBELPMDLSEKSEPOE
EGSPAGGOGKGAPAEAGGCAETLPASAATENCCCTPKERGAAPAGCEOKENTLES
MKMGSRRLRGJGWESSLRORPMRLTQPDADPYIATSKRDEWLAMKEAEAKKA
KVIGANNAVENOGQGVVEESLPVAQQPTDPASPATTPPEPVGSDDAQGNATK
AGDGEPEYEDGFGIGELWVGKLGKGFSGMWPGRIVSMWTGSRRAABETRMYWMFDDG
KFSVCYVKMLPSPCSAFHQAIVYNQCPMRAIYEVLQVASRAGKRFPPVCHDBE
SDTKAVENVONKMPIEMALGGFOFGPGKLEPEBEENKPIKEYVTDTMWPEAAAYAP
PPPAKRPKSTAEKPVKYKEIIDERTRELRYEVROKRNIEDICISGSLNLTLEHL
FVGGCAQAIKEDPNMCYGDCKGTITGLLRBDEMPSRLQMFNNNDGDFEPKYVP
VPGGAQAAILKEDPNMCYGDCKGTITGLLRBDEMPSRLQMFNNNDGDFEPKYVP
PPPAKRPKIRIVLSFDGIALGLVLDLGIQDVRTIASVECDSTIVGMVRHQGKM
YVDVRSKTRIVOLHIOGPFDLVIGSGPNDLSINPARKIYEGTGILFEFTRLHD
ARPKGCDNRPPFMLPENVALVGVGDKRDISRFLSNPNIMDAKESAAHFARYFWGL
PEMNPLASTVNDDLEOCLRHRIAKFSKRVITITRSSNIKGGKHOFPPVNEKE
DLTWCTERNRVGFPHVITTVDSNNRLARKQLLGRSVSPVIRHLFAPLKEYFACV"

ORIGIN

Alignment Scores:

| | Pred. No.: | Length: | Score: |
|-----|------------------------|---------------|---------|
| QY | 0 | 3005 | 3763.00 |
| Db | 3763.00 | Matches: | 689 |
| | Percent Similarity: | Conservative: | 100.0% |
| | Best Local Similarity: | Mismatch: | 0 |
| | Query Match: | Indels: | 100.0% |
| DB: | Gaps: | 0 | 5 |

US-10-623-813-86 (1-689) x AP067972 (1-3005)

QY 1 MetasnlavajluciuasnnglnlyprvgiyugiserGIntLyvaIgUjuaLasr 20
Db 899 ATGAATGCTGTGGAAAAACAAGGGGCCCGGGAGACTTCAGAGAAGTTGAGAGGCACG 958
QY 21 ProProhlaVaJaIngInPrOThraApProLaaserProThValAlaThrThProGu 40
Db 959 CCTCTCTCTGTGACGACGCCCACTGACCCTCCGCGCATCCCCACGTGTGTAACACCGCTGAG 1018

| | Accession | Gene | Protein | Length | |
|------------|---|---|---------|--------|-----------------|
| Db | 2059 | CCACCTGTCGCCAGCTGAGAAAGGAAAGCCCATCCGGGAGCTGCTCTCTTGAATGGAATC | | 2158 | |
| Qy | 4421 | AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTryIleAlaSer | | 440 | |
| Db | 2159 | GCTACAGGGGCTCTGTGCTGTAAGACATTGGGCATTCAAGTGGACCGGTAACATTGGCTCG | | 2218 | |
| Qy | 441 | GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTry | | 460 | |
| Db | 2219 | GAGGTGTGTGAAGCATCTCCATCAACGGTGGGCATGTGGCGGCACCAAGGGGAAGTCAATGAC | | 2278 | |
| Qy | 461 | ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTryProPheAspLeu | | 480 | |
| Db | 2279 | GTCGGGGAAGCTCCGACGCTCAACAAGACATATTCAGAGTGGGGCCCATTCATCTCG | | 2338 | |
| Qy | 481 | ValIleGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu | | 500 | |
| Db | 2339 | GTGATTTGGGGGAGTCCCTGCAATGACCTTCATGTCATGACCCCTGCTCGAAGGGCTTC | | 2398 | |
| Qy | 501 | TryGlnGlyThrGlyArgLeuPhePheGlnPheTryArgLeuLeuHisAspAlaArgPro | | 520 | |
| Db | 2399 | TACAGAGGACATGGCCGGCTCTTCTTGAATCTACCCGCTCCGTCATATGGCGGCC | | 2458 | |
| Qy | 521 | LysGlnGlyAspAspArgProPhePheTrpLeuPheGlnAsnValValAlaMetGlyVal | | 540 | |
| Db | 2459 | AAGGAGGAGATATATGCCCCCTTCTTGAGCTTTTGAAGATGTGGGGCCATGGGGCTT | | 2518 | |
| Qy | 541 | SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys | | 560 | |
| Db | 2519 | AGTACAAAGGGACATCTCGCATTTCTTCGAGTCCACCTGTGATGATTGATGCCAA | | 2578 | |
| Qy | 561 | GluValSerAlaAlaHisArgAlaArgTryPheTrpGlyAsnLeuProGlyMetAsnArg | | 580 | |
| Db | 2579 | GAAGTGTGAGTCAACAACAGGGCCCGTACTTCTGGGGTAACTTCCCGGTATGAACAG | | 2638 | |
| Qy | 581 | ProLeuAlaSerThrValAsnAspLysLeuGlnLeuGlnLysCysLeuGlnHisGlyArg | | 600 | |
| Db | 2639 | CCGTGGCAATCCACTGTGATGATGAAGCTGGAGCTGAGAGATGTCTGGAGCATGGCAG | | 2698 | |
| Qy | 601 | IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly | | 620 | |
| Db | 2699 | ATGAGCAAGTTCAGCAAGTGAAGACATTACTACAGAGTCAAACTCCATTAAGCAAGGC | | 2758 | |
| Qy | 621 | LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTrpCysThrGlu | | 640 | |
| Db | 2759 | AAAGACACGACATTTTCTGCTTCATGAATGAGAAAGAGACATCTTATGGTGACAGAA | | 2818 | |
| Qy | 641 | MetGluArgValPheGlyPheProValHisTryThrAspValSerAsnMetSerArgLeu | | 660 | |
| Db | 2819 | ATGAAAAGGGATATTGGTTTCCAGTCCACTATACGACGTCTCCAACTAGAGCCGCTTG | | 2878 | |
| Qy | 661 | AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla | | 680 | |
| Db | 2879 | GCGAGGACAGACTGCTGGGCGGCTCATGAGCTGCGCATGCAATCCGCCACTTTCGCT | | 2938 | |
| Qy | 681 | ProLeuLysGlyTryPheAlaCysVal | | 689 | |
| Db | 2939 | CCGCTGAAGAGTATTTTGGCTGTG 2965 | | | |
| RESULT 4 | CS050242 | 3604 bp | DNA | linear | PAT 23-MAR-2005 |
| LOCUS | CS050242 | | | | |
| DEFINITION | Sequence 26 from Patent WO2005021757. | | | | |
| ACCESSION | CS050242 | | | | |
| VERSION | CS050242.1 | GI:61889497 | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | Homo sapiens (human) | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; | | | | |
| | Homnidae; Homo. | | | | |
| REFERENCE | 1 | | | | |
| AUTHORS | Tomme, P.H. and van Rompaey, L. | | | | |

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OW protein - nucleic search, using frame_p2n model

Run on: October 5, 2006, 18:12:34 ; Search time 10519 Seconds
(without alignments)
6282.884 Million cell updates/sec

Title: US-10-623-813-86
Perfect score: 3763
Sequence: 1 MNAAVENOGSGESQKVEAS.....WSVPVIRHLFAPLKEYFACV 689

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 segs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DRV=xlh
-Q=/abs/ABSSWEB.spool/US10623813/runat_05102006_121146_27659/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10623813 @CN_1_1_6706 @runat_05102006_121146_27659 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -XGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

GenEmbl: *
1: gb_env: *
2: gb_pat: *
3: gb_ph: *
4: gb_pl: *
5: gb_pr: *
6: gb_ro: *
7: gb_sts: *
8: gb_sy: *
9: gb_un: *
10: gb_vl: *
11: gb_ov: *
12: gb_hcg: *
13: gb_in: *
14: gb_cm: *
15: gb_ba: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 3763 | 100.0 | 2371 | 5 | AF480163 | AF480163 Homo sapi |
| 2 | 3763 | 100.0 | 3005 | 2 | CS050243 | CS050243 Sequence |
| 3 | 3763 | 100.0 | 3005 | 5 | AF067972 | AF067972 Homo sapi |

| | | | | | | |
|----|--------|-------|--------|----|----------|--------------------|
| 4 | 3763 | 100.0 | 3604 | 2 | CS050242 | CS050242 Sequence |
| 5 | 3763 | 100.0 | 4258 | 5 | AF331856 | AF331856 Homo sapi |
| 6 | 3763 | 100.0 | 4294 | 2 | DD171352 | DD171352 Marker fo |
| 7 | 3763 | 100.0 | 4294 | 5 | BC043617 | BC043617 Homo sapi |
| 8 | 3763 | 100.0 | 4395 | 2 | CS050241 | CS050241 Sequence |
| 9 | 3763 | 99.5 | 2841 | 2 | C0720469 | C0720469 Sequence |
| 10 | 3714 | 98.7 | 2318 | 6 | AF480164 | AF480164 |
| 11 | 3714 | 98.7 | 4094 | 6 | BC007466 | BC007466 Mus muscu |
| 12 | 3714 | 98.7 | 4152 | 6 | AF068625 | AF068625 Mus muscu |
| 13 | 3712 | 98.6 | 2597 | 14 | AY271299 | AY271299 Bos tauru |
| 14 | 3702 | 98.4 | 4476 | 5 | AB208833 | AB208833 Homo sapi |
| 15 | 3570.5 | 94.9 | 3014 | 11 | AB214886 | AB214886 Gallus ga |
| 16 | 3570.5 | 85.7 | 264380 | 12 | AC120824 | AC120824 Rattus no |
| 17 | 3225 | 82.0 | 253474 | 12 | AC095441 | AC095441 Rattus no |
| 18 | 3085.5 | 81.7 | 3377 | 11 | AB196917 | AB196917 Danio rer |
| 19 | 3076 | 80.4 | 235547 | 12 | AC112586 | AC112586 Rattus no |
| 20 | 3024.5 | 78.2 | 2933 | 11 | AB196919 | AB196919 Danio rer |
| 21 | 2944 | 65.0 | 344787 | 12 | AC112040 | AC112040 Rattus no |
| 22 | 2447.5 | 64.1 | 2191 | 5 | AK025230 | AK025230 Homo sapi |
| 23 | 2413 | 60.9 | 3544 | 11 | AB196915 | AB196915 Danio rer |
| 24 | 2292.5 | 58.6 | 168651 | 5 | AC009474 | AC009474 Homo sapi |
| 25 | 2204.5 | 57.8 | 4135 | 6 | AF068627 | AF068627 Mus muscu |
| 26 | 2174.5 | 57.7 | 4163 | 6 | AF151974 | AF151974 Mus muscu |
| 27 | 2172.5 | 57.7 | 4278 | 6 | AF151970 | AF151970 Mus muscu |
| 28 | 2172.5 | 57.7 | 4278 | 6 | AF068626 | AF068626 Mus muscu |
| 29 | 2169.5 | 57.7 | 4195 | 6 | AF151973 | AF151973 Mus muscu |
| 30 | 2167.5 | 57.6 | 4223 | 6 | AF151973 | AF151973 Mus muscu |
| 31 | 2167.5 | 57.6 | 4338 | 6 | AF151969 | AF151969 Mus muscu |
| 32 | 2165 | 57.5 | 4267 | 5 | AF176228 | AF176228 Homo sapi |
| 33 | 2159 | 57.4 | 3630 | 6 | BC105922 | BC105922 Mus muscu |
| 34 | 2157 | 57.3 | 4043 | 2 | C0726303 | C0726303 Sequence |
| 35 | 2157 | 57.3 | 4445 | 5 | AF156488 | AF156488 Homo sapi |
| 36 | 2157 | 57.3 | 4335 | 2 | CQ981515 | CQ981515 Sequence |
| 37 | 2157 | 57.3 | 4335 | 5 | AF331857 | AF331857 Homo sapi |
| 38 | 2149.5 | 57.1 | 4369 | 5 | DO321787 | DO321787 Homo sapi |
| 39 | 2062 | 54.8 | 2491 | 14 | AY244710 | AY244710 Bos tauru |
| 40 | 2012.5 | 53.5 | 6248 | 11 | AB196918 | AB196918 Homo rer |
| 41 | 1948.5 | 51.8 | 2421 | 14 | AY244712 | AY244712 Bos tauru |
| 42 | 1902 | 50.5 | 3746 | 6 | AF068628 | AF068628 Mus muscu |
| 43 | 1900 | 50.5 | 3344 | 6 | AF151976 | AF151976 Mus muscu |
| 44 | 1900 | 50.5 | 4089 | 6 | AF151971 | AF151971 Mus muscu |
| 45 | 1897 | 50.4 | 4006 | 6 | AY078427 | AY078427 Mus muscu |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|------|--------|-----------------|
| RESULT 1 | AF480163 | 2371 bp | mRNA | linear | PRI 14-OCT-2002 |
| LOCUS | AF480163 | | | | |
| DEFINITION | Homo sapiens DNA cytosine methyltransferase 3A2 (DNMT3A2) mRNA, | | | | |
| ACCESSION | AF480163 | | | | |
| VERSION | AF480163.1 | GI:23954437 | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | Homo sapiens (human) | | | | |
| REFERENCE | Homo sapiens | | | | |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| TITLE | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; | | | | |
| | Homnidae; Homo. | | | | |
| | 1 (bases 1 to 2371) | | | | |
| | Chen,T., Ueda,Y., Xie,S. and Li,E. | | | | |
| | A Novel Dnmt3a Isoform Produced from an Alternative Promoter | | | | |
| | Localizes to Euchromatin and Its Expression Correlates with Active | | | | |
| | de Novo Methylation | | | | |
| JOURNAL | J. Biol. Chem. 277 (41), 38746-38754 (2002) | | | | |
| PUBMED | 12138111 | | | | |
| REFERENCE | 2 (bases 1 to 2371) | | | | |
| AUTHORS | Chen,T. and Li,E. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (31-JAN-2002) Cardiovascular Research Center, | | | | |
| | Massachusetts General Hospital, 149 13th Street, Charlestown, MA | | | | |
| | 02129, USA | | | | |
| FEATURES | Location/Qualifiers | | | | |

```
source
1..2371
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p23"
1..2371
/gene="DNMT3A2"
217..2286
/function="de novo DNA cytosine methylation"
/codon_start=1
/product="DNA cytosine methyltransferase 3A2"
/protein_id="AA040037.1"
/db_xref="gi:23954438"
/translation="MNAVEENQGEESQKEEDASPPVAQOPTDPASPTVAITTPPEVGS
DAGKNATKAGDDPEYEDGRGFGIGELVWKLRFSPWRIVSMMWTGSRAEGT
RWVMFGDFKSVVCEVKLMLPSFCSAFHQATYKQPMYKALYEVLQVASSRAGKL
FPVCHDSDESDTAKAEVONKPMIEMALGFQSPGKLEPPEEKNIKEYTIDMNV
EPBAAYAPPPPAKKPKRKSTAEKPKVKEIIIDERTBERLVYEROKCRNIEDICSGS
LNTLLEHPLFYGCKONCKNCFLECAVYDDDGQSYCTICCGREVMCGMNNCCRC
FCVCEVDLVGPGAQAQAIKEDPMVCYCKHGTYGLRKREDMPSRLQMFPANNHDQ
EFDPKXYPPVPABKRKPIRVLSLFDIATGLVLKDIGIVDRIASEVEDSIITVG
MVRHQKIMYGDVRSVTOKAIQEWGPDLYGSPNDLSIVNPARKGLVEGTGRLF
FEFVRLHDARKEGDDRPFLPENVVAMGVSDKRDISRELSNPVMIIDAKESAAH
PARYFMKNLPGMNRPLASTVNDKLEQCELEHGRILAKSKYRRTITTSNSIKQKDOH
PVPVMEKEDILMCTEMERVRGFPVHYTIDVSNMSRLARQLLGMSKSVPIRLHAPL
KEYFACV"

Alignment Scores:
Score: 0 Length: 2371
Percent Similarity: 3763.00 Matches: 689
Best Local Similarity: 100.0% Conservative: 0
Query Match: 100.0% Mismatches: 0
Indels: 0
Gaps: 0

US-10-623-813-86 (1-689) x AF480163 (1-2371)

QY 1 MetAsna1aValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
Db 217 ATGATGCTGTGGAAGAAACCGAGGGCCCGGGAGTCTCAGAAAGTCGAGAGGCCAGC 276
QY 21 ProProAlaValGlnGlnInProThrAspProAlaSerProThrValAlaThrThProGlu 40
Db 277 CCTCCTGCTGTGACGACGCCACTGACCCCGCATCCCTCCTGCTGCTACCAAGCTTGA 336
QY 41 ProValGlySerAspAlaGlyAspLysAsna1aThrLysAlaGlyAspAspGluProGlu 60
Db 337 CCCGTGGGCTCGATGCTGGGGACAAAGATGCCACCAAGACAGCGAGACGACCAAG 396
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTTPGlyLysLeuArgGlyPhe 80
Db 397 TACGAGGACGGCCGGGCTTGGCACTGGGAGCGGTGTGGGGAAACTCGCGGGGCTTC 456
QY 81 SerTTPTrpProGlyArgIleValSerTTPTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db 457 TCCGTGTGGCCAGCGCGCATGTGTCTTGTGTGATGAGACGGCGCGGACCGAGCACTGAA 516
QY 101 GlyThrArgTTPValMetTTPPheGlyAspGlyLysPheSerValCysValGluLys 120
Db 517 GGCACCCCGCTGGGTCACTGTTCGGAGACCGCAAAATCTCAGTGTGTGTGTGAGAG 576
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db 577 CTGATGCCGCTGAGACTCGTTTTCAGTGTGTTCACACGAGCCACGATCAACAAGCAGCCC 636
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db 637 ATGTACCGCAAAAGCAATCTACGAGGTCTGTGACAGGTGCACAGCGCGGCGGAGACTGG 696
QY 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValAsn 180
```

```
Db 697 TTCCGGTGTGCCACGACGAGGATGAGATGACCTCCAAAGCCGTGAGGTCCAGAAC 756
QY 181 LysProMetIleGluTTPAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db 757 AAGCCCATGATTGAATATGGCGCTTGGGGGCTTCCAGCCCTTCTGCGCCCTTAAGCGCTGAG 816
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTPValGluPro 220
Db 817 CCACCAAGAGAGAGAAATCCCTTACAAAGATGTACAGGACATGTGGTGGAACTT 876
QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db 877 GAGCAGACTGCTTACGACACCTCCACGACCAAAAGCCCCGGAAGAGACAGCCGGJ 936
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db 937 AAGCCCAAGGTCAAGAGATTAATGATGAGCGCAAGAGAGCGGCTGTGTACGAGGT 996
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValTyr 280
Db 997 CGGCAAGTGTCCGGAACTTGAAGACATCTGCATCTCTGTGGAGCTTCAATGTTACC 1056
QY 281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db 1057 CTGGAACACCCCTCTTCTGTGAGGAATGTGCCAAACTGCAAGAACTGCTTCTGAG 1116
QY 301 CysAlaTyrGlnTyrAspAspArgLysTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1117 TGTGCTTACCAAGTACGACGACGAGGCTTACGATCTGACCATCTGCTGTGGGGG 1176
QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1177 GGTAGGTGCTCATGTGTGGAAACAACTGCTGCAAGTCTTTCGTGAGATGTGTG 1226
QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaIleLysGluAspProTyrAsnCys 360
Db 1237 GACCTTGGTGGGGCCGGGGGCTGCCAGGACCATTAAGAAACCCCTGGAATGTC 1296
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTTPProSer 380
Db 1297 TACATGTGCGGCAAGAGGTACTTACGGGCTGTGGCGGCGGCAAGAGACTGGCCCTCC 1356
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProLysValTyr 400
Db 1357 CGGCTTCAGATGTTCTTCTGCTAATACCAAGACAGGAAATTGACCTCCAAAGTTTAC 1416
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 1417 CCACTGTCCCAAGCTGAGAAAGAGAACCCATCCGGGTGCTGTCTCTTTGATGGAATC 1476
QY 421 AlaThrGlyLeuLeuValIleLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 1477 GCTACAGAGGCTCTCGGTGTGAAGACTTGAGCAATTAAGTGAAGCCCTCATTTGCCCTCG 1536
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 1537 GAGGTGTGTAGACTCTCATCACCGTGGCATGTGGGACCAAGGGAAGATGATGTATAC 1596
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTTPGlyProPheAspLeu 480
Db 1597 GTCGGGAGACTCCCGACGTCACACAGAAACATATCCAGAGTGTGGGCCCATTTGATCTTG 1656
QY 481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db 1657 GTGATTTGGGGAGGTCTCTCAATGACTCTTCAATCTCTCAACCTGTCTCGCAAGGACTTC 1716
QY 501 TyrGlnGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 1717 TACAGGGCACTGGCCGGCTCTTCTTGAATTCACCGCTCTCGCATGATGCGCGGCC 1776
QY 521 LysGluGlyAspAspArgProPhePheTyrLeuPheGluAsnValValAlaMetGlyVal 540
```

Db 1777 AAGAGAGATGATGATGCCCCCTTCTTGCGCTTTTGAGATGTGTGGCCATGGCGTT 1836
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 1837 AGTGAACAGAGGAGCATCTCGCATTTCTCGATCCCAACCTGTGATGATGCCAA 1896
QY 561 GluValSerAlaIleHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db 1897 GAAGTGTCAAGCTCACACAGGCCCCCTACTTCTGGGGTAACTTCCGGTATGAACAGG 1956
QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnLysCysLeuGlnHisGlyArg 600
Db 1957 CCGTTGGCATCCCACTGTGAATGATTAAGCTGGAGCTGAGGAGGTCTGTGGAGCATGGCAGG 2016
QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db 2017 ATAGCCCAAGTTCAGCAAGAGGAGGACCATTAACGAGTCAAACTCCATTAAGACAGGCG 2076
QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGlnAspIleLeuTyrCysThrGlu 640
Db 2077 AAGACACACATTTCTCTGTCTTCATGAATGAGAAAGAGACATCTTATGTGCACCTGAA 2136
QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2137 ATGGAAAGGATATTGTGTTCCAGTCCACTATACGACCTTCCAACATGAGCCGCTTG 2196
QY 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2197 GCGAGGCAAGAGACTGCTGGGCGCGTCATGGAGCGTGCACAGTCAACGCCACTTTCGCT 2256
QY 681 ProLeuLysGluTyrPheAlaCysVal 689
Db 2257 CCGTGAAGAGATATTTCGCTGTGTG 2283

RESULT 2
LOCUS CS050243 3005 bp DNA linear PAT 23-MAR-2005
DEFINITION Sequence 27 from Patent WO2005021757.
ACCESSION CS050243
VERSION CS050243.1 GI:61889498
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1.
REFERENCE
AUTHORS Tome, P.H. and van Rompaey, L.
TITLE Polypeptides and polynucleotides for use as a medicament
JOURNAL Patent: WO 2005021757-A 27 10-MAR-2005;
Galapagos Genomics N.V. (BE)
FEATURES
source
1.3005
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3005
Score: 3763.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-623-813-86 (1-689) x CS050243 (1-3005)

QY 1 MetAsnAlaValAlaGluLysGlnGlyProGlyGluSerGlnLysValAlaGluAlaSer 20
Db 899 ATGAATGCTGTGGAGAAACCAAGGGGCCCGGGAGAGTCTCAGAAAGTGTGGAGGCCAGC 958
QY 21 ProProAlaValAlaGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40

Db 959 CCTCCGTGCTGTGACAGAGCCCACTGACCCCCGCACTCCCACTGTGGCTACCAAGCCTGAG 1018
QY 41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
Db 1019 CCGTGTGGGTCCGATGCTGGGACAAGAAATGCCACCAAGCAGGCGCATGACAGCCAGAG 1078
QY 61 TyrGluAspGlyValArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuAspGlyPhe 80
Db 1079 TAGAGAGAGGCGCGGGCTTTGGCATGGGGAGCTGTGTGGGGAACCTGCCGGCTTC 1138
QY 81 SerTrpTrpProGlyValArgIleValSerTrpTrpMetThrGlyValSerArgAlaAlaGlu 100
Db 1139 TCTGTGGCCAGAGCCGCACTTGCTTGCTTGATGATACGGGCCGAGCCGAGCAGCTGAA 1198
QY 101 GlyThrArgTyrValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
Db 1199 GGCACCCGCTGGGTCAATGTGTTCGGAGACGGCAAAATCTCAGTGCTGTGTGGAGAG 1258
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db 1259 CTGATGCTGCTGAGCTCGTTTGTCACTGCTTCACACAGGCCACGTAAACAAGCAGCCC 1318
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db 1319 ATGTACCGCAAGCCATCTACAGAGTCTGCAAGTGGCCAGAGCCGCGGGGGAAGCTG 1378
QY 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValAlaGlnAsn 180
Db 1379 TTCCTGGTGTGCACAGCAGCATGAGTGAACATGCCAAGGCCGTGGAGGTGCAGAAC 1438
QY 181 LysProMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db 1439 AAGCCCATATTGAATGAGGCCCTGGGGGCTTCCACCTTCTGGCCCTTAAGGGCCTGGAG 1498
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db 1499 CCACGAGAAAGAGAGAAATCCCTACAAAGAAAGTACACCGACATGTGGGTGAACCT 1558
QY 221 GluAlaAlaIleTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db 1559 GAGGAGCTGCTACGACACCTCCACCGCCAAACACCCCGGAAAGACAGCGGGAG 1618
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db 1619 AAGCCCAAGGTCAAGGAGATTTATGATGAGCGCACAGAGAGCGCTGTGTACGAGGTG 1678
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 1679 CCGCAGAAATGGCGGAACATTGAGGACATCTGCATCTCTGTGGAGCCCAATGTTACC 1738
QY 281 LeuGlnHisProLeuPheValGlyLysMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db 1739 CTGGAACACCCCTCTCTGTGGAGGAATGTGCCAAATGTCAAGAACTGCTTTCGGAG 1798
QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1799 TGTGCTACAGAGCAGCAGCGAGCTTACAGTCTTACTGACCATCTGTCTGTGGGGC 1858
QY 321 ArgGluValLeuMetCysGlyLysAsnAsnAsnCysCysArgCysPheCysValGlyCysVal 340
Db 1859 CGTGAGTGTCTCATGTGTGGGAAACAAACATCTGTGAGAGTGTGGTGGAGGTGTG 1918
QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaIleLysGluAspProTyrAsnCys 360
Db 1919 GACCTCTTGTGTGGGCGCGGGCTGTGCCAGGACCATTAAGGAACCCCTGGAACCTGC 1978
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLysLeuLeuArgArgArgLysAspTrpProSer 380
Db 1979 TACATGTGGGGCAAGAGTACCTTACGGGCTGTCTGTGGCGGGAAGAGATGCGCCTCC 2038
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400

| | | | | | |
|------------|--|--|------|--------|-----------------|
| Db | 203 | CGGCTCAGATGTTCTTCGCTAATACCAACGACGAAATTGACCTCCAAAGTTTAC | 2098 | | |
| Qy | 401 | ProProVal1ProAlaGluIuYarGlyProIleArgValLeuSerLeuPheAspGlyIle | 420 | | |
| Db | 2099 | CCACCTGTCCACGCTGAGAGAGAGAACCCATCCGGGTGCTGTCTCTTGTATGAGAAATC | 2158 | | |
| Qy | 421 | AlaThrGlyLeuLeuValLeuIuYasPheGlyIleGlnValAspArgTyrIleIleAspSer | 440 | | |
| Db | 2159 | GCTACAGGGCTCTCTGTGCTGAAGACTTGGGCATTCAGGTGACCGGCTACATTCCTCTCG | 2218 | | |
| Qy | 441 | GluValCysGluAspSerIleThrValGlyMetCValArgHisGlnGlyLysIleMetTyr | 460 | | |
| Db | 2219 | GAGGTGTGTGAGCATTCATCACGGTGGCCAGTGGCCGACACAGGGAGATCATTTGAC | 2278 | | |
| Qy | 461 | ValGlyAspValArgSerValThrGlnIuYasHisIleGlnGlyTyrGlyProPheAspLeu | 480 | | |
| Db | 2279 | GTCGGGGACGTCGCCAGCGTCAACAGAAAGCATATCCAGAGTGGGGCCCATTCGATCTG | 2338 | | |
| Qy | 481 | ValIleGlyGlySerProCysAsnAspLeuSerIleValAspProAlaArgLysGlyLeu | 500 | | |
| Db | 2339 | GTGATTGGGGGACGCTCTGCAATGACCTTCCTCATGTCACCTGCTCGCAAGGGCTTC | 2398 | | |
| Qy | 501 | TyTGluGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro | 520 | | |
| Db | 2399 | TACGAGGCACCTGGCGGCTCTTCTTTGAGTTCTACCGCTCTCTCATGATCGCGGCC | 2458 | | |
| Qy | 521 | LysGluGlyAspAspArgProPhePheTyrLeuPheGluAsnValAlaIleMetGlyVal | 540 | | |
| Db | 2459 | AAGGAGGAGATGATCGCCCTTCTTCTGCGCTCTTGAAGATGTGGTGGCCATGGGCGTT | 2518 | | |
| Qy | 541 | SerAspIuYasArgAspIleSerArgPheLeuGluIuSerAsnProValMetIleAspAlaLys | 560 | | |
| Db | 2519 | AGTACCAAGAGGAGCATCTCGGATTTCTCGAGTCCAACTCGATATATGATGCCAA | 2578 | | |
| Qy | 561 | GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg | 580 | | |
| Db | 2579 | GAAGTGTACGTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCGGTATGAACAGG | 2638 | | |
| Qy | 581 | ProLeuAlaSerThrValAsnAspIuYasLeuGluLeuGlnCysValLeuGluHisGlyArg | 600 | | |
| Db | 2639 | CCGTTGGCATCCTACGTGAATCATAGCTGAGCTGACGAGAGTGTCTGAGCATGGCAGG | 2698 | | |
| Qy | 601 | IleAlaIuYasPheSerLeuValArgThrIleThrThrArgSerAsnSerIleLysGlnGly | 620 | | |
| Db | 2699 | ATAGCAAGTTCAGCAAGTGAAGTGAACCTTACTACGAGCTCAACTCCATTAAGCAGGACC | 2758 | | |
| Qy | 621 | LysAspGlnHisPheProValPheMetAsnGluYasGluAspIleLeuTyrCysThrGlu | 640 | | |
| Db | 2759 | AAAGACCAACATTTTCTGTCTTCTTGAATGAAAGAGACATTTATGGTGCACTGAA | 2818 | | |
| Qy | 641 | MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu | 660 | | |
| Db | 2819 | ATGGAAGAGGTATTTGGTTCCCACTCCACTATCATGACGCTCCCAACATGAGCGGCTTG | 2878 | | |
| Qy | 661 | AlaArgGluArgLeuGluArgSerTrpSerValProValIleArgHisLeuPheAla | 680 | | |
| Db | 2879 | GCGAGGCGAAGACGCTGGCGGCTCATGAGCGGTCCAGATCATCCGCACTTTCCT | 2938 | | |
| Qy | 681 | ProLeuIuYasGluTyrPheAlaCysVal | 689 | | |
| Db | 2939 | CCGCTGAAGAGATTTTGGCGTGTG | 2965 | | |
| RESULT 3 | AF067972 | 3005 bp | mRNA | linear | PRI 12-FEB-2001 |
| LOCUS | AF067972 | | | | |
| DEFINITION | Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA, complete cds. | | | | |
| ACCESSION | AF067972 | | | | |
| VERSION | AF067972.2 GI:12746531 | | | | |
| KEYWORDS | Homo sapiens (human) | | | | |
| SOURCE | Homo sapiens | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |

| | | | |
|--|---|--|------|
| REFERENCE | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo. | | |
| AUTHORS | 1 (bases 1 to 3005) Xie,S., Wang,Z., Okano,M., Nogami,M., Li,Y., He,W.W., Okumura,K. and Li,E. | | |
| TITLE | Cloning, expression and chromosome locations of the human DNMT3 gene family | | |
| JOURNAL | Gene 236 (1), 87-95 (1999) | | |
| PUBMED | 10433969 | | |
| REFERENCE | 2 (bases 1 to 3005) Xie,S., Okano,M. and Li,E. | | |
| AUTHORS | Submitted (25-MAY-1998) CVRC, Mass. General Hospital, 149 13th Street, Charlestown, MA 02129, USA | | |
| TITLE | 3 (bases 1 to 3005) Hata,K., Shirohzu,H., Sasaki,H. and En,L. | | |
| JOURNAL | Direct Submission Submitted (12-FEB-2001) CVRC, Mass. General Hospital, 149 13th Street, Charlestown, MA 02129, USA | | |
| REMARK | Sequence update by submitter On Feb 12, 2001 this sequence version replaced gi:4927369. | | |
| COMMENT | Location/Qualifiers | | |
| FEATURES | 1..3005 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="2" /map="2p23; between D2S171 and D2S174; 48-50 cM" 1..3005 /gene="DNMT3A" 230..2968 /gene="DNMT3A" /function="de novo DNA cytosine methyltransferase" /product="DNA cytosine methyltransferase 3 alpha" /protein_id="AAD3084.2" /db_xref="GI:12746532" /translation="MPAMSSGPGPDPTSSAAAREEDRKDGEDEBPGRKEBROESTT ARKVRGRGRKRHPVESGDTTPKDPATVYSKSPMAOAGSSELIPNDLEKRSERPE EESPAGGKGGAAPABEGGALETLPASRAVENGCCTKEGRGALEAGKEOKENTIS MKREGSGRGLRGWESSLRORPMPLTFOADPYITSRKDEMLARMKEXKA KVIAGMAVENOQPGESQVEASPPAVOQPDPAATTPPEVGSADAGKNATK AGDDEPEYEDGREGIGELVWKLRCGSPMGKIVSWMTGRSALAGTVMVMFGD KFSVYEXKMLPISFCSAFOATVYNQPMYRARIYVLOVASRAKLFPVCHDSDE SDTRKAVYVONKPMIEMALGROPBGPKLEPEPEENKPKVAYTDMVPEAAAYAP PPPAKPKRSTAEKPKVETIDERTRELYEVRQKRNLEDICISGSLNTLEHL FVGMCQNCNCELECAVYQYDDGYOSTYCTICCGREVLKCGNNCCRCFCVBCVDL VCPGAQAAPAIKEDPMNOCMGHGTGLRLRBDMPRLQMFANNHDOEDPDKVP PVAERKRPPIRVLSLFDGIATGLVLDLGIQVDRYIASEVCEDSITVGMVHOGKIM YVGDVRSVTKHIOEMGPFDLVIGGSPCNDLSIVNPAKGLYEGLGLFEFYRLHD APKREDDRPFPMLFERNVVMGYSDDKRDLSRPLESNPVMIDAKESVAARARVWGL PGNMRPLASTVNDKLELOECLERGLAKESKVTITTRRSNIKQGDQHPVPFNNEKE DILMCTEMERVFGFPVHYTDVSNMSRLARQLLRGSVSPVIRHLPALKEFYACV" | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 0 | | |
| Score: | 3763.00 | | |
| Percent Similarity: | 100.0% | | |
| Best Local Similarity: | 100.0% | | |
| Query Match: | 100.0% | | |
| DB: | 5 | | |
| US-10-623-813-86 (1-689) x AF067972 (1-3005) | | | |
| QY | 1 | MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnYasValGluGluAlaSer | 20 |
| DB | 899 | ATGAATCCTGTGGAAACAAACAGGGGCCCGGGGAGTCTCAGAAAGGTGAGAGCCAGC | 958 |
| QY | 21 | ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrTrpProGlu | 40 |
| DB | 959 | CCTCTCTGTGACGAGCCCACTGACCCCGCATCTCCCACTGTGCTACCAACCCCTGAG | 1018 |

QY 41 ProValGlySerAspAlaGlyAspLeuAspAlaThrValAlaGlyAspAspGluProGlu 60
 DB 1019 CCGGTGGGGTCCGATGCTGGGGACAGAAATGCCACCAAGCAGCGATGACGACCCAGAG 1078
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
 DB 1079 TAGAGAGACGGCGGGCTTTGGCATTGGGGAGCTGGGTGGGGAAATCGCGGGCTTC 1138
 QY 81 SerTrpTrpProGlyValGlyIleValSerTrpTrpMetThrGlyValSerAlaAlaGlu 100
 DB 1139 TCTGTGGGGCCAGGCCCATTTGTCTTGTGGATGACGGGCCGAGCCAGCAGCTGAA 1198
 QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 1199 GGCACCCGGTGGGATGATGGGTGGAGACGGCAAAATTCACGTGGTGTGTGGAGAG 1258
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGluPro 140
 DB 1259 CTGATGCCGTGAGCTGTTTTCAGTGGCTCCACAGGCCACGTCACACAGCAGCCC 1318
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerAlaGlyLysLeu 160
 DB 1319 ATGTACCGCAAGCCATCTACAGGTCTCGAGGTGCCAGAGCCGGCGGGAACCTG 1378
 QY 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValAsn 180
 DB 1379 TTCGGGTGGCCACGACGACGATGACGTGACCTCCAGGCCGTGGAGGTGACAGAC 1438
 QY 181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 DB 1439 AACCCATGATTTGAATGGGCTGGGGGCTTCAGCCCTTGCCCTAAGGGCTGGAG 1498
 QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
 DB 1499 CCACCGAAGAAAGAAAGAAATCCCTCAAGAGTGTACACGCAATGTGGTGAACCT 1558
 QY 221 GluAlaAlaIleAlaTyrAlaProProProAlaLysLysProAlaGlyLysSerThrAlaGlu 240
 DB 1559 GAGGACAGCTGCTACGACCACTCCACAGCCAAAGGCCCGGAGAGACAGCCGAG 1618
 QY 241 LysProLysValLysGluIleIleAspGluArgThrArgLysLeuValTyrGluVal 260
 DB 1619 AACCCAGGTCAAGAGGATTTATGATGAGGACAGAGAGCGGCTGGTACGAGTG 1678
 QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1679 CGGCAGAGGTCCGGAACTTGAGGACATCTGCATCTCTGTGGAGCCCAATGTAC 1738
 QY 281 LeuGluHisProLeuPheValGlyLysMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 DB 1739 CTGGAAACCCCTCTTGTGGAGATGTGCCAAACCTGCAAGACTGTTTCTGAG 1798
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 DB 1799 TGTGGCAGACAGATCGACGAGAGGCTACAGTCCACTGACACATCTGTGTGGGGC 1858
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 DB 1859 CGTGAGGTCTCATGTGCGGAAACCAACACTGCTGCGAGTCTTTCTGTGAGTGTG 1918
 QY 341 AspLeuLeuValGlyProGlyValAlaAlaGluAlaIleLysGluAspProTyrAsnCys 360
 DB 1919 GACCTCTGGTGGGGCGGGGGCTGCCAGGCAACCTTAAGGAAGCCCTGGAATGCG 1978
 QY 361 TyrMetCysGlyLysLysGlyThrTyrGlyLeuLeuArgArgArgLysAspTrpProSer 380
 DB 1979 TACATGTGGCGGACAGAGGATCTTACGGCTGTGTGGGGCGCAGAGGACTGGCCCTCC 2038
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 DB 2039 CGGCTCCAGATGTTCTTCGCTAAATTAACAGACACAGCAATTTGACCTCCAAAGTTTAC 2098
 QY 401 ProProValProAlaGlyLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420

DB 2099 CCACTGTCCCACTAGAAAGAGAGGCCATCCGGGTGTGTCTCTTGTGATGGAATC 2158
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 DB 2159 GCTACAGGGCTCCTGTGTGTGAAGGACTTGGGCAATTCAGGTGACCGCTACATTCCTCG 2218
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2219 GAGGTGTGAGGACTTCATACCGTGGCATGTGTGGGACCAAGGGAAGATCATTTAC 2278
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
 DB 2279 GTGCGGGAGCTCCGACGCTCACAGAGCATATCCAGAGTGGGGCCCATTCGATCTG 2338
 QY 481 ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2339 GTGATTTGGGGCAGTCCCTGCATATGACCTTCATGTCAACCTGCTGCCAAGGGCTC 2398
 QY 501 TyrGluGlyThrGlyValArgLeuPhePheGluPheTyrArgLeuHisAspAlaArgPro 520
 DB 2399 TACGAGGACCTGGCCGGCTCTTTCAGTTCTACGCTCTTCATGATGTCGGCGGCC 2458
 QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
 DB 2459 AAGAGGAGATGATGCGCCCTTCTTCTGCTCTTGAGAAATGTGTGGCCATGGCGCTT 2518
 QY 541 SerAspLysArgAspIleSerArgPheLeuGlnSerAsnProValMetIleAspAlaLys 560
 DB 2519 AGTGACAGAGGAGACATCTCGCATATTCGAGTCCAAACCTGTGATGATGATGCCAA 2578
 QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyValAsnLeuProGlyMetAsnArg 580
 DB 2579 GAAGTGTACCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCGGGTATGAACAG 2638
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnLysCysLeuGluHisGlyArg 600
 DB 2639 CGTTGGCATCCACTGATGATGATTAAGCTGAGCTGACGAGAGTGTGGAGATGCGAGG 2698
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 DB 2699 ATAGCCAGTTCAGCAAGTGAAGACCATTTACATACAGGTCAACCTCCATAAGCAGGGC 2758
 QY 621 LysAspGluHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
 DB 2759 AAAGACCAAGATTTCTGTCTTCAATGATGAAAGAGACATCTTATGTGCATCGAA 2818
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2819 ATGGAAGGATATTGTGTTCCAGTCCACTATACGACGTCCAAACATGAGCGGCTTG 2878
 QY 661 AlaArgGlnArgLeuLeuGlyValGserTrpSerValProValIleArgHisLeuPheAla 680
 DB 2879 GCAGAGCAGAGACTGTGGCCCGTCATGAGCGTCCCACTCATCGCCACCTTCGCT 2938
 QY 681 ProLeuLysGluTyrPheAlaCysVal 689
 DB 2939 CCGCTGAAGAGTATTTGCGTGTGTG 2965
 RESULT 4
 CS050242 3604 bp DNA linear PAT 23-MAR-2005
 LOCUS CS050242
 DEFINITION Sequence 26 from Patent WO2005021757.
 ACCESSION CS050242
 VERSION CS050242.1 GI:61889497
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1
 AUTHORS Tomme, P. H. and van Rompaey, L.

TITLE polypeptides and polynucleotides for use as a medicament
JOURNAL Patent: WO 2005021757-A 26-10-MAR-2005;
Galapagos Genomics N.V.(BE)
FEATURES
source Location/Qualifiers
1..3604
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3604
Score: 3763.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-623-813-86 (1-689) x CS050242 (1-3604)
QY 1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
DB 217 ATGAATCTCTGGAGAAACCAAGGGCCCGGGAGTCTCAGAAAGTCGAGAGCCAGC 276
QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
DB 277 CTTCTGCTGTGGAGCAAGCCACTGACCCGCACTCCCACTGTGGTACCAAGCCCTGAG 336
QY 41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
DB 337 CCGGTGGGTCCGATGCTGGGGACAAGATGCCAACAGCAGCGATGACGAGCCAGAG 396
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
DB 397 TACGAGACCGCCCGGCTTTGGCATTGGGAGCTGTGTGGGAAACTGCGGGGCTTC 456
QY 81 SerTrpTrpProGlyArgLysIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
DB 457 TCTGTGGTGGCAAGGCCGCTATTGTCTTGGTGTGATGACGGGCCGAGACCGAGCTGAA 516
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
DB 517 GGCACCCGCTGGGTCATGTGGTTCGAGACCGCAAAATCTCAGTGTGTGTGAGAAAG 576
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheIleGlnAlaThrTyraAsnLysGluPro 140
DB 577 CTGATGCCCTGACCTCGTTTGGAGTGCCTTCCACAGGCCACGTAACAACAGCAGCCC 636
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 637 ATGATCCGCAAGCCATCTACAGAGTCTCTGACAGTGGCCAGCACCCGCGGGAACTCG 696
QY 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
DB 697 TTCCCGGTGTGCCACGACAGGATGAGTGACACTGCCAAGGCCGTGAGGTGCAGAAC 756
QY 181 LysProMetIleGluTyrPalaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 757 AAGGCCATGATTGAATGGCCCTGGGGGCTTCCAGGCTTCTGGCCCTTAAGGGCTGGAG 816
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
DB 817 CCACGAAAGAAAGAAAGAAATCTCTTACAAAGAAATGTACAGGACATGTGGTGGAACT 876
QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
DB 877 GAGGCACTGCTACGACCACTCCACCAAGCAAAAAGCCCCGGAAGAGACAGCCGAG 936
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgLysGluLeuValTyrGluVal 260
DB 937 AAGCCCAAGGTCAAGGAGATTATGATGACGCAACAAGAGCGGCTGTGTACGAGGTG 996
QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280

DB 997 CGGCAAGAGCCGGAAACATTGAGGACATCTGCATCTCTGTGGAGCCCAATGTTACC 1056
QY 281 LeuGlnHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 1057 CTGGAACACCCCTCTTCTGTGGAGAAATGTGCCAAAATGCAAGAACTCTTCTTGAG 1116
QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
DB 1117 TGTGGTACAGTACGACGACGAGCTTACCATCTTACACCATCTCTGTGGGGCC 1176
QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
DB 1177 GGTGAGGTCTCATGTGGCGGAAACAACTGCTGACGGTGTCTTTCGTGAGTGTGTG 1236
QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTyrPheCys 360
DB 1237 GACCTCTTGTGTGGGCGGGGGCTGGCCACGAGCCATTAAAGAAAGCCCTGGAACTGC 1296
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGluAspTrpProSer 380
DB 1297 TACATGTGGGACAAAGGTACTTACGGGCTGCTGGCGGAGAGACTGGCCCTCC 1356
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
DB 1357 CGGCTCAGATGTTCTTCGCTTAATAACAGACAGCAAGAAATTGACCTCCAAAGTTTAC 1416
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 1417 CCACCTGTCCCACTGAGAAAGAAAGCCCATCCGGGTGTGCTCTTGTGATGAAATC 1476
QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
DB 1477 GCTACAGAGGCTCTGTGTCTGTAAGGACTTGGGCATTCAGGTGACCGCTCATTCCTCG 1536
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 1537 GAGGTGTGAGAGATCCATCACGATGGCATGTGCGGACCAAGGAAAGATCATGTAC 1596
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
DB 1597 GTGGGAGACTCCGACGGCTCACAGAACATATCCAGAGAGGGGCCATTCGATCTG 1656
QY 481 ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 1657 GTGATTTGGGGACGTCTCTGCATGACTCTTCATGTCAACCTGTGCAAGGGCTTC 1716
QY 501 TyrGluGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
DB 1717 TACGAGGGCACTGGCGGCTCTTCTTGAATCTTACCGCTCTGCAATGATGGCGGCCCC 1776
QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
DB 1777 AAGAGGGAATGATAGCCCTTCTTGTGCTTTGAGATGTGGGCAATGGAGGGGTT 1836
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
DB 1837 AGTGACAAAGGGAACATCTGCGATTTCTCGAGTCCAACCTGTGATGATGCCAAA 1896
QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
DB 1897 GAAGTGCAGCTCAACAAGGGCCCGCTACTTCTGGGGTAACTTCCCGGTATGAACAG 1956
QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnLysCysLeuGlnHisGlyArg 600
DB 1957 CCGTTGGCATCCCTGTGAATGATTAAGCTGAGCTGAGAGAGTGTGAGCATGGCAGG 2016
QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
DB 2017 ATAGCCAAAGTTCAAGAAAGTGAAGACATTACTACGAGGTCAAACTCATTAAGCAGGGC 2076
QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrPyrThrGlu 640

| | | | |
|------------|------------|--|--------|
| Db | 2077 | AAAGACGACATTTTCCTGCTTCATGATGAGAAAGAGCATCTTATGTCAGCTGAA | 2136 |
| Qy | 641 | MetGluArgValPheGlyPheProValHisTyrThrSerPylSerAsnMetSerArgLeu | 660 |
| Db | 2137 | ATGGAAAGGGTATTTGGTTTCCAGTCCACTATACACGCTCCCAACATGAGCGGCTTG | 2196 |
| Qy | 661 | AlaArgGlnArgLeuGlyArgSerTyrSerValProValIleArgHisLeuPheAla | 680 |
| Db | 2197 | GCGAGCGAGAGACTGCTGGGCGGCTCATGAGCGTGCACATGCCCTCTTGCT | 2256 |
| Qy | 681 | ProLeuLysGluTyrPheAlaCysVal | 689 |
| Db | 2257 | CCGCTGAAGAGTATTTGCTGTGTG | 2283 |
| RESULT 5 | AF331856 | 4258 bp | linear |
| LOCUS | AF331856 | Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA, | |
| DEFINITION | AF331856 | complete cds. | |
| ACCESSION | AF331856.1 | GI:18033252 | |
| VERSION | AF331856.1 | GI:18033252 | |
| KEYWORDS | | Homo sapiens (human) | |
| SOURCE | | Homo sapiens | |
| ORGANISM | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo. | |
| REFERENCE | | 1 (bases 1 to 4258) | |
| AUTHORS | | Ni,J., Pradhan,S. and Roberts,R.J. | |
| JOURNAL | | Cloning, expression and characterization of human DNMT3 genes | |
| AUTHORS | | 2 (bases 1 to 4258) | |
| TITLE | | Unpublished | |
| REFERENCE | | Ni,J., Pradhan,S. and Roberts,R.J. | |
| JOURNAL | | Direct Submission | |
| AUTHORS | | Submitted (22-DEC-2000) New England Biolabs, 32 Tozer Road, | |
| TITLE | | Beverly, MA 01915, USA | |
| FEATURES | | Location/Qualifiers | |
| source | | 1..4258 | |
| | | /organism="Homo sapiens" | |
| | | /mol_type="mRNA" | |
| | | /db_xref="taxon:9606" | |
| | | /chromosome="2" | |
| | | 1..4258 | |
| | | /gene="DNMT3A" | |
| | | 227..2956 | |
| | | /gene="DNMT3A" | |
| | | /codon_start=1 | |
| | | /product="DNA cytosine methyltransferase 3 alpha" | |
| | | /protein_id="AA57039.1" | |
| | | /db_xref="GI:18033253" | |
| | | /translation="MSSSGRGDTSSSAAREEDRKQGEBOEPGRKEERPEPTARK VGPGRKRKHPIVESGDTKPDPAVISKSPWADSGASBELPVGDELEKSEPEPESG PAGGCGGAGAPAEAGEAETLPEASRAVENCCPTPKGRGAPAGAEQKNTNESKM EGSRGGLRGAGESSLRQRPMLRTFOADPYIISKRKDEWLAKREAEKAVI AGNAAVENNGPESQKEASPPAOOPDPAFTVAATPPEADGADKATKAD DEPEYDGGGFGTGLVWGKLRSFSPWPGIIVSMVTGRRABEGRWMMFGDGKFS VVCVEKLPLSFSFCSAFOHTYTKQEMTKRATLIEVQVASRAGKLPRVCHDSDPT AKAVEQNKPEIMWALJGFQPSGPKGRLPEEEKNPKEYITDMWPEEAATAPPP AKKPRSTAEKPKVKEIIDERTRERLVEYRQKRNIEDICISGSLNLTLEHPLVG GAKQKCNKFLFECAYDDGYOYSTICCGGREVLNMGNNCCRFCEVCVLLVGP GMAQQAIXEDPNMCIWCGHKGTGLRLRRRDMPSRLQMFPAHHNDEFDPKYPPV AERKRIRLQSLFEDGIATGLLVLKDIGIOWDRYIASVECDSTTVGMVRGQKIMVG DRSVSTQKHTIQEWGPFDDLVYGSPCNDLSVNPARKGLYEGSTRALFFERRLIHADP KEGDDPFPFMLENVNVMGVSDKDISRFLESNPVMTDAEVAARHARTFMGNLGM NRPPLSTVNDKLEQCLEHGRITAKFSKVTITTRSNSIKQGDQHFPFVFNKEKITL WCTEMERVFGFPVHYTDVSNMSRLARQLLGRSWSVPVIRHLFAPLKEFYACV" | |
| gene | | | |
| CDS | | | |

Alignment Scores:
 Pted. No.:
 Score:
 Percent Similarity:
 Best Local Similarity:

0
 3763.00
 100.0%
 100.0%

Length:
 Matches:
 Conservative:
 Mismatches:

4258
 689
 0
 0

| | | | |
|--|--------|---|------|
| Query Match: | 100.0% | Indels: | 0 |
| DB: | 5 | Gaps: | 0 |
| US-10-623-813-86 (1-689) x AF331856 (1-4258) | | | |
| Qy | 1 | MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer | 20 |
| Db | 887 | ATGAATGCTGTGAAGAAAACAGAGGGCCCGGGAGCTCAGAAAGGTGAGAGGCCAGC | 946 |
| Qy | 21 | ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu | 40 |
| Db | 947 | CCCTCGCTGTGCGAGCGCCACTGACCCCGCATCCCACTGTGTACAGCCCTGAG | 1006 |
| Qy | 41 | ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu | 60 |
| Db | 1007 | CCCGTGGGTCCATGCTGGGACAGAAATGCCACCAAGCAGGCTGACGAGCCAGAG | 1066 |
| Qy | 61 | TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe | 80 |
| Db | 1067 | TACGAGGACGGCCGGGCTTTGGCATTTGGGAGCTGTGTGGGGAAACTGCGGGCTTC | 1126 |
| Qy | 81 | SerTyrTrpProGlyArgGlyLeuValSerTyrTrpMetThrGlyArgSerArgAlaAlaGlu | 100 |
| Db | 1127 | TCTGTGGTGGCAGCGCCGCAATGTGTCTGTGTGATGACGGCCGGAGCCGAGAGCTGAA | 1186 |
| Qy | 101 | GlyThrArgTyrValMetTyrPheGlyLysPylLysPheSerValValCysValGluLys | 120 |
| Db | 1187 | GACACCCGCTGGGTCTATGTGTTCGAGACGCGCAATTCATGTGTGTGTGAGAG | 1246 |
| Qy | 121 | LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro | 140 |
| Db | 1247 | CTGATCCGCTGAGCTCGATTTCAGTGTGCTTCACAGGCCAGCAACAGCAGCC | 1306 |
| Qy | 141 | MetTyrArgGlySerAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu | 160 |
| Db | 1307 | ATGTACCGCAAAACCACTTACGAGGTCTGCAAGTGGCCAGCGCGCGGAGACTG | 1366 |
| Qy | 161 | PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluAlaAsn | 180 |
| Db | 1367 | TTCCCGGTGTGCACACACACATGAGATGACATGCCCAAGCCGTGAGGTGACAGAC | 1426 |
| Qy | 181 | LysProMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu | 200 |
| Db | 1427 | AAAGCCATGATTAATGAGGCTGGGGGGGCTTCACGCTTCTGCGCTTACAGGCTGAG | 1486 |
| Qy | 201 | ProProGluGluGluLysAsnProTyrLysGluValTyrThrSpmectTyrValGluPro | 220 |
| Db | 1487 | CCACCGAAGAAAGAAAGATCCCTCAAGAAAGTGTACCGGACATGTGGTGGAACTT | 1546 |
| Qy | 221 | GluAlaAlaAlaTyrAlaProProProAlaLysLysProArgLysSerThrAlaGlu | 240 |
| Db | 1547 | GAGGAGCTGCTTACGACACACCTCCACAGCCCAAAAAGCCCGGAAGACACAGCGAG | 1606 |
| Qy | 241 | LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal | 260 |
| Db | 1607 | AAAGCCAGGTCAAGAGATTATGTATGAGCGGCAAGAAGAGCGGCTGTGTACAGGTG | 1666 |
| Qy | 261 | ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr | 280 |
| Db | 1667 | CGGCAAAATGCGCGGAACATTGAGCATCTGATCTCTGTGGAGGCTCAATGTTAAC | 1726 |
| Qy | 281 | LeuGlnHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnLysPheLeuGlu | 300 |
| Db | 1727 | CTGGAAACACCCCTCTTCTGTGGAGAAATGTGCCAAAATCGCAAGAACTGCTTTCGAG | 1786 |
| Qy | 301 | CysAlaIleTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly | 320 |
| Db | 1787 | TGTGCGTACCAAGACGACGAGCGGCTACAGTCTTACCTGCTGTGTGGGGGC | 1846 |
| Qy | 321 | ArgGluValLeuMetCysGlyLysAsnAsnLysCysArgCysPheCysValGluCysVal | 340 |
| Db | 1847 | CGTGAAGTCTCATGTGCGAAACAAACACTGTCGAGGTGCTTTGCGGTGAGTGTG | 1906 |

| | | | |
|----|------|--|------|
| QY | 341 | LeuLeuValGlyProGlyValAlaGlnAlaAlaIleValGluAspProThrPheGly | 360 |
| Db | 1907 | GACCTCTTGAGGGGCGGGGGGCTCGGACGAGCCATTAAAGAAAGACCTCGGAACCTGC | 1966 |
| QY | 361 | TyrMetCysGlyPheIleGlyGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer | 380 |
| Db | 1967 | TACATCTCGGGGCACAAGGGTACCTACGGGCTCTCTCGGGCGGAGAGACCTGGCCCTTC | 2022 |
| QY | 381 | ArgLeuGlnMetPhePheValAsnAsnHisAspGlnGluPheAspProProLeuValTyr | 400 |
| Db | 2027 | CGGCTCCAGATGTCTTCGCTAATAAACCAACGACACAGAAATTTGACCTCTCAAGGTTTAC | 2086 |
| QY | 401 | ProProValProAlaGluIleAsnArgLysProIleAlaGlyValLeuSerLeuPheAspGlyLe | 420 |
| Db | 2087 | CCACCTGTCACAGCTGAGAGAGAGAGCCATCCGGGTGCTGTCTCTTGATGGAAATC | 2146 |
| QY | 421 | AlaTrpGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer | 440 |
| Db | 2147 | GCTACAGGAGCTCTGTGTGTGAAGACTTTGGGCATTCAAGTGGACCGCTACATTCTCTCG | 2206 |
| QY | 441 | GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr | 460 |
| Db | 2207 | GAGGTGTGTAGAGACTCCATCACTGGTGGGACATGTGGGACACAGGGAGATCATGTAC | 2266 |
| QY | 461 | ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu | 480 |
| Db | 2267 | GTCGGGGAGCTCCGACGGCTCAACAGAAAGCATATCCAGAGATGGGGCCATTTCATCTGTG | 2322 |
| QY | 481 | ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu | 500 |
| Db | 2327 | GTGATTTGGGGGAGAGCTCCCTGCACATACCTCTCATCTGTACACCTTGCTCGCAAGGACCTC | 2386 |
| QY | 501 | TyrGlnGlyThrGlyValArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro | 520 |
| Db | 2387 | TACAGAGGCCACTGGCGGCTCTCTTTAGATTCTAACCGGCTCTGCATGATCGGGGGCCC | 2446 |
| QY | 521 | LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal | 540 |
| Db | 2447 | AAGGAGGAGATGATCGGCCCTCTTCTTGCGCTCTTGAGAAATGTGTGGCCATGGGCGCTT | 2506 |
| QY | 541 | SerAspLysAlaArgAspIleSerArgPheLeuGlnGlySerAsnProValMetIleAspAlaLys | 560 |
| Db | 2507 | AGTGCACAAAGGAGACATCTCGGATTTCTCGAGTCCAAACCTGTGATGATGATGATCCCAA | 2566 |
| QY | 561 | GluValSerAlaAlaHisAsrGlyAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg | 580 |
| Db | 2567 | GAGGTCTCAGCTGCACACAGGGCCCGCTACTCTCTGGGGTAACCTTCCGGTATGAACAGG | 2622 |
| QY | 581 | ProLeuAlaSerThrValAsnAspLysLeuGlnGluCysLeuGluHisGlyArg | 600 |
| Db | 2627 | CCGTTGGCATCCACTGTGATGATGATGAAGCTGAGCTGCAGAGATGTCTTGAGCATGGCAGG | 2686 |
| QY | 601 | IleAlaLysPheSerLysValArgHisIleThrTrpArgSerAsnSerIleLysGlnGly | 620 |
| Db | 2687 | ATAGCCAAAGTTCAGCAAAAGTGAGGCCCATTTACTACAGAGTCAACTCATTAAGAGAGGAGC | 2746 |
| QY | 621 | LysAspGlnHisPheProValPheMetAsnGlnLysGluAspIleLeuTrpCysThrGlu | 640 |
| Db | 2747 | AAAGCCACGATTTTCTCTCTCTTCAATGAATGAGAAAGAGACATTTATGGTGCCTGA | 2806 |
| QY | 641 | MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu | 660 |
| Db | 2807 | ATGGAAAGGATATTGTGTTTCCCAATCCACTATACTGACGTCTCCAAATGAGCGCGCTTG | 2866 |
| QY | 661 | AlaArgGlnArgLeuLeuGlyValArgSerTrpSerValProValIleArgHisIleuPheAla | 680 |
| Db | 2867 | GCGAGGCGAAGACTGTCTGGCCGGCTCATGAGACGCTGCCAGTCATCTGCCACCTCTTCCCT | 2922 |
| QY | 681 | ProLeuLysGluTyrPheAlaCysVal | 689 |
| Db | 2927 | CCGCTGAAGAGATATTTGCGGTGTG | 2953 |

[illegible]

Oy 161 PheProValCysHisAspSerAspGluSerAspThrAlaValAlaGluValGlnAsn 180
 Db 1387 TTCGGGCTGTCACAGCAGCAGTATGAGTACACTGCGCAAGCGCTGAGAGTCAGAAC 1446
 Oy 181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProGlyGlyLeuGlu 200
 Db 1447 AAGCCCATGATTGAATGGCCCTGGGGGGCTTCCAGCCTTCTGGGCCCTTAAGGCCCTGAG 1506
 Oy 201 ProProGluGluGluValAsnProTyrIleGluValTyrThrAspMetTyrValGluPro 220
 Db 1507 CCACCAAGAGAGAGAAATCCCTCAAAAGATGTACAGCAGCATGTGGGTGAACCT 1566
 Oy 221 GluAlaAlaAlaTyrAlaProProProProAlaValIleValSerProAlaValSerThrAlaGlu 240
 Db 1567 GAGGCACTGCTTACGACCACTCCACAGCAAAAGCCCGGAGAGACAGCAGCGAG 1626
 Oy 241 LysProLysValIleValGluIleIleAspGluValArgThrArgIleValTyrGluVal 260
 Db 1627 AAGCCCAAGCTCAAGAGATTTATGATGAGCGCACAGAGAGCGGCTGGTACAGAGTG 1686
 Oy 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 Db 1687 CGGCAGAGATGCCGGAACATTGAGGACATCTGCACTCTCTGGGAGGCTCAATGTATACC 1746
 Oy 281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnValSerValSerPheLeuGlu 300
 Db 1747 CTGGAAACACCCCTCTTCTGTGGAGAAATGTCCAAACCTGCAAACTGCTTTCTGGAG 1806
 Oy 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 Db 1807 TGTGCGGACAGTACGACAGAGAGCGCTACCACTCTTACCACTCTGCTGTGGGGGC 1866
 Oy 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 Db 1867 CGTGAGGTGCTCACTGTCGGAACCAACACTGTCAGAGTCTTTGCGTGAAGTGTGTG 1926
 Oy 341 AspLeuLeuValGlyProGluValAlaAlaGlnAlaIleValGluAspProTyrAsnCys 360
 Db 1927 GACCTCTTGTGGGGCCGGGGGGCTGCCAGCACCATTAAAGAGAACCCCTGTGAACGTC 1986
 Oy 361 TyrMetCysGlyHisIleValTyrTyrGlyLeuLeuValArgArgGluAspTyrProSer 380
 Db 1987 TACATGTGCGGGCCCAAGAGGTAACCTTACGGGCTGTGGCGCGCAGAGAGACTGCCCTCC 2046
 Oy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 Db 2047 CGGCTCCAGATGTTCTTCGTAATACACAGCAGAGAAATTGACCTCCCAAGGTTTAC 2106
 Oy 401 ProProValProAlaGluValArgValProIleArgValLeuSerLeuPheAspGlyIle 420
 Db 2107 CCACCTGTCTCCAGCTGAGAGAGAGAACCCCATCGGGGTGCTCTCTTGTGATGAGATC 2166
 Oy 421 AlaThrGlyLeuLeuValIleLeuValAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 Db 2167 GCTACAGGGCTCTCTGCTGCTGAGAGACTTGGGCATTCAGGTGACCCCTCATTTGCTCG 2226
 Oy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 Db 2227 GAGGTGTGTAGAGCATCATCAACGCTGGGCTGTGGCGCACCGAGGAAAGATCATGTAC 2286
 Oy 461 ValGluValAspValArgSerValThrGlnLysHisIleGlnGluTyrProPheAspLeu 480
 Db 2287 GTGGGGGACCTCCGACAGCTCACCAAGATATCCAGGAGTGGGGCCCATTTGATCTTG 2346
 Oy 481 ValIleGlyGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 Db 2347 GTGATTTGGGGGCACTCCCTGCAAGACTCTCCACTGCTCAACCTGCTCCAGAGGCTC 2406
 Oy 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 Db 2407 TACGAGGGCACTGGCGGCTCTTCTTGTGATTAACGCTCTCGATGATGCGGGCCC 2466

Oy 521 LysGluValAspAspArgProPheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
 Db 2467 AAGAGGAGAGTGAATCCCTCTTCTTGTGCTTTTAAAGATGTGGGGCCATGGCGTT 2526
 Oy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
 Db 2527 AGTACAGAGAGGAGCATCTCGCATTTCTCGAGTCCAAACCTGTGATGATGATGCCAA 2586
 Oy 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyValLeuPheProGlyMetAsnArg 580
 Db 2587 GAACTGTCACTGTGACACAGAGCGCGCTTCTTGGGGTAACTTCCCGGTATGAACAGG 2646
 Oy 581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGluCysLeuGluHisGlyArg 600
 Db 2647 CCGTTGGCATTCACCTGTGAATGATTAAGCTGAGCTGCAGAGATGTCTGAGACATGGCAGG 2706
 Oy 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 Db 2707 ATAGCCAAAGTTCAAGCAAGTGAAGACCATTAACAAGAGTCAAACTCATTAAGCAGGGC 2766
 Oy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
 Db 2767 AAAGACCAAGCATTTTCTGTCTTCATGATGAGAAAGAGACATCTTATGTGTCACCTGA 2826
 Oy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 Db 2827 ATGGAAAGGATTTGTTGTTTCCAGTTCATTAATGACGTTCTTCAACATGAGCGCTTG 2886
 Oy 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
 Db 2887 GCGAGCAGAGACCTGCTGGCGCGGTCAAGAGCGTGCAGTCAATCCGCCACTTTCGCT 2946
 Oy 681 ProLeuLysGluTyrPheAlaCysVal 689
 Db 2947 CCGCTGAAGAGATATTTGCGTGTGTG 2973
 RESULT 7
 LOCUS BC043617
 DEFINITION Homo sapiens DNA (cytosine-5'-methyltransferase 3 alpha, transcript variant 3, mRNA (cDNA clone MGC:50948 IMAGE:6150112), complete cds.
 ACCESSION BC043617
 VERSION BC043617.1 GI:27694443
 KEYWORDS MCC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 4294)
 Stransberg,R.L., Feigold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shemmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Millan,S.J., Bosnak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.R., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Schevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932

REFERENCE 2 (bases 1 to 4294)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: "http://mgc.nci.nih.gov
REMARK COMMENT
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Maason, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prambu, Parvaneh Saeedi, Jr Santos, Angelique Scherch, Ursula Skaleka, Duane Smalton, Jeff Stoct, Miranda Teai, George Yang, Jacque Schein, Aaim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAC Plate: 89 Row: 0 Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

SOURCE

1. 4294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:50948 IMAGE:6150112"
/cissue_type="Skin, melanotic melanoma."
/clone_lib="NIH MGC_72"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 4294
/gene="DNMT3A"
/note="Synonyms: DNMT3A2, M.Hsallita"
/db_xref="GeneID:1788"
/db_xref="MIM:602769"
238. 2976
/gene="DNMT3A"
/codon_start=1
/product="DNA cytosine methyltransferase 3 alpha, isoform a"

CD5
gene
protein_id="AAH43617.1"
/db_xref="GI:2769444"
/db_xref="GeneID:1788"
/db_xref="MIM:602769"
/translation="MPAMPSGPGDPTSSAAREDRKDGEGEORPEKREORBSPT
ARKYGRPKRKHPPVESGDTKCPAVISKPSMAODGASELLPNEGLEKRESEQPE
EGSAGGQKGAPAGPABEGAAETLPEASAVENGCCPREGGAPEAGKEKEKTES
MKMGSGRRLRGGLGWESSLRORPWRJTFQAGDPYYSKREDEWLARWRREKKA
KVIAGMAVENOGFESQKVEASPPVQOQTPASPTVAATPEPVSGDAGDNATK
AGDEPEYEDRGFGIGELVWGKLRGFMWGRIVSMWTRSRRAEGTRVMWMDG
KESVYCEKMLPSFCSAFOATYNNQPMYRKAIYELVQASSRAGLTPVCHDSDE
SDTKAAVEONKPMIENALGFGFQSGPGLPEPEEKPYKEVTTDNVNEEAAAYAP
PPPAKPKRSTAEKPKVKEIITDERRELVIEVRKCNITDICTSCSLVNTLEHLP
FVGAGCONCNCFLECAQYDDQYOSCTICCGREVLWGNANNCCGCFVECDVL
VGPAAQAAIKEDPWNCYMGHKGTYLKRREDPWSRLQWFPANNHDOEDPPKYV
VPVPEKKRPIVLSPDGIATGLVLKQIGQVDRYIASECEDSITGVAVHQQKLV
YVCGVSTQKILQEMGEPDLVIGSSPNDISVNPARKGLYEGTGRLPFRFVLLHD
ARPEGDRPFPMLEPENVVAMGVSDKDRISFLSNPMTDAKEYSAAHRRARYWGNL
PGNMRPLASTVNDKLELQECLEHGRIAKFSVRITITTSNSIKQGDHPFVEMEXS
DILWCTEMERVFEPVHYTDVSNMRLRRLRLGRSMSPVIRHLFAPLKEYFACV"

ORIGIN

Alignment Scores:

| Pred. No.: | 0 | Length: | 4294 |
|------------------------|---------|---------------|------|
| Score: | 3763.00 | Matches: | 689 |
| Percent Similarity: | 100.0% | Conservative: | 0 |
| Best Local Similarity: | 100.0% | Mismatches: | 0 |
| Query Match: | 100.0% | Indels: | 0 |
| DB: | 5 | Gaps: | 0 |

US-10-623-813-86 (1-689) x BC043617 (1-4294)

| | | | |
|----|------|---|------|
| QY | 1 | MetasnlalavalgiugluasnginglyProclYgluSerglnlyvalgluGluAlaSer | 20 |
| DB | 907 | ATGAATGCTGTGAAGAAACACAGGGCCCGGGAGTCTCAGAGGTGAGAGGCCAGC | 966 |
| QY | 21 | ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThProGlu | 40 |
| DB | 967 | CCTCTGCTGTGCGACGCCACCTAGACCCCGCATCCCATCTGGCTACACAGCCTGAG | 1026 |
| QY | 41 | ProValGlySerAspAlaGlyAspIlyAsnAlaThrIlysaIagIyAspAspGluProGlu | 60 |
| DB | 1027 | CCCGTGGGCTCGATGCTGGGACAGAAATGCCACAAAGCAGCGATGACGAGCCAGG | 1086 |
| QY | 61 | TyrlGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyIlySleuArgIlyPhe | 80 |
| DB | 1087 | TACGAGGACGGCCGGGCTTTGGCATTGGGAGCTGGTGTGGGAAACTCGGGGCTTC | 1146 |
| QY | 81 | SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaIleGlu | 100 |
| DB | 1147 | TCCGATGGCCAGGCCCATTTGTCTCTTGATGATGAGGAGGAGGCGGAGCGGAGCTGAA | 1206 |
| QY | 101 | GlyThrArgTrpValMetTrpPheGlyAspGlyIlyPheSerValValCysValGluIlyS | 120 |
| DB | 1207 | GGCACCGCTGGGTCATGTGCTTGAGAGCGGAAATTCATGAGTGTCTGTTGAGAG | 1266 |
| QY | 121 | LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrIlyAsnIlySglnPro | 140 |
| DB | 1267 | CTGATGGCGGTGAGCTGCTTTTTCAGTGGCTTCACAGGCGCATCAACAGAGCCG | 1326 |
| QY | 141 | MetIlyrArgIlysaIaIeIyTrglValLeuGlnAlaIaSerSerArgAlaGlyIlySleu | 160 |
| DB | 1327 | ATGTACCGCAAGCACTACAGAGTCTGTGAGTGGCCAGACCCGCGGGAAGCTG | 1386 |
| QY | 161 | PheProValCysHisAspSerAspGlySerAspThrAlaIlysaIaValGluValGlnAsn | 180 |
| DB | 1387 | TTCCCGGTGTCCACGACGAGATGACATCTCCAAAGCCGTGTGAGTCTCAGAAC | 1446 |
| QY | 181 | IlySProMetIleGluTrpAlaLeuGlyIlyPheGlnProSergIlyProIySgIlyLeuGlu | 200 |
| DB | 1447 | AAGCCATGATTGAATGGCCCTTGGGGGCTTCCAGCTTCTGGCCCTTAAGGGCCTGAG | 1506 |
| QY | 201 | ProProGluGluGluIlyAsnProIlyIlyGluValIlyThrAspMetTrpValGluPro | 220 |
| DB | 1507 | CCACAGAAAGAGAAATCCCTCAAGAAATGTACAGGACATGTGGTGAACCT | 1566 |
| QY | 221 | GluAlaIaIaIyTrAlaPProProProAlaIlySlySProAlGlySserThAlaGlu | 240 |
| DB | 1367 | GAGCACTGCTTACGACACCTCCACACGCAAAAGCCCGGAGAGACAGCCGAG | 1626 |
| QY | 241 | IlySProIlysaIlySgIlyIleIleAspGluArgThrArgIluArgLeuValIyTrglVal | 260 |
| DB | 1627 | AAGCCAAAGGCAAGGAGATTATGATGAGGCAACAAGAGCGCGTGTGTACAGAGTG | 1686 |
| QY | 261 | ArgGlnIlySlysArgaAnIleGluAspIleCysIleSergCysIlySserLeuAsnValThr | 280 |
| DB | 1687 | CGGAGAGAGTCCGGAACATTGAGACATCTGCTCTGTGGAGCTCAATGTTTACC | 1746 |
| QY | 281 | LeuGluIlySProLeuPheValIglIlyMetCysGlnAsnCysIlyAsnCysPheLeuGlu | 300 |
| DB | 1747 | CTGGAACACCCCTCTTCTGTGAGAGATGTGCCAAAGCAAGCAAGCACTGTTCTGGAG | 1806 |
| QY | 301 | CysAlaIyTrGlnIyAspAspAspGlyIyTrGlnSerIyCysThrIleCysCysGlyGly | 320 |
| DB | 1807 | TGTGCTTACCACTACGACGACGAGCTTACAGCTTACCACTGTGCTGTGGGGC | 1866 |

[illegible][illegible]

QY 181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProGlyLeuGlu 200
 DB 1548 AAGCCCATGATTGAATGGGCCCTGGGGGCTTCCAGCTTCTGGCCCTAAAGGGCTGGAG 1607
 QY 201 ProProGluGluGluValAsnProTyrIleGluValTyrThrAspMetTrpValGluPro 220
 DB 1608 CCACCAAGAAAGAAAGAAATCCCTCAAAAGAAAGTACAGGACATGTGGGTGAACTT 1667
 QY 221 GluAlaIa 240
 DB 1668 GAGGACACTGCTTACGACCACTCCACCAACCAAAAGCCCGGAAGAGCAGCGGAG 1727
 QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
 DB 1728 AAGCCCAAGGCTCAAGAGATTAATGATGAGCGCAACAGAGCCGCTGGTGTACAGAGTG 1787
 QY 261 ArgGlnLysCysAspAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1788 CCGCAAGAGTCCCGAACATTGAGACATCTGCATCTCTGTGGAGCCTCAATGTTACC 1847
 QY 281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 DB 1848 CTGGAAACACCCCTCTTCTGTTGAGAGATGTGCCAAAACCTCAAGAACTGCTTCTGGAG 1907
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 DB 1908 TGTGCTTACCACTACGACGACGACGACGCTTACCTACTGACCATCTGCTGGGGGGC 1967
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 DB 1968 CGTAGGGCTCACTGTGCGAAACAACTGCTGACAGTGTCTTTGGCTGAGTGTGTG 2027
 QY 341 AspleuLeuValGlyProGlyValAlaGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIa 360
 DB 2028 GACCTCTTGGTGGGGGCGGGGGCTGCCAGCAGCATTAAAGAAACCTCGGAATGCG 2087
 QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgGluAspTrpProSer 380
 DB 2088 TACATGTGCGGGCAACAGGTAACCTACCGGCTGTGCGGGCGGAGGACTGGCCCTCC 2147
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 DB 2148 CGGCTCCAGATGTTCTTGCTTAATACACACACGAAATTGACCTTCCAAAGTTTAC 2207
 QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 DB 2208 CCACCTGTCCAGCTGAGAAAGAGGACCCATCCGGGTGCTGCTCTTTGATGGAATC 2267
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
 DB 2268 GCTACAGGGCTCTGTGCTGTAAGCATTTGGGCAATTCAGGTGAGCCGCTAATGCTCCG 2327
 QY 441 GluValCysGluAspSerIleThrValGlyMetValAspHisGlnGlyLysIleMetTyr 460
 DB 2328 GAGGTGTGTGAGACTCATCAACGTTGGGCAATGTGCGGACACAGGGAAGATCATGTAC 2387
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
 DB 2388 GTGGGGGACGTCCGAGCGTCACACAGAGCATATCCAGAGTGGGGCCCATTCATCTG 2447
 QY 481 ValIleGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2448 GTGATTTGGGGCAAGTCTGTGCAATGACTTCCATCTTCAACCTGTGCGAGGGGCTTC 2507
 QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 DB 2508 TACAGGGCACTGGCGGCTCTTCTTAAGTTCTACCCGCTCTGCAAGATGCGGGGCC 2567
 QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
 DB 2568 AAGAGGAGATGATGCGCCCTTCTTGTGGCTCTTGAAGAAATGTGTGGCCATGCGGCTT 2627

QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
 DB 2628 AGTACAAAGGGGACATCTGCCGATTTCTGAGTCCAACTGTGATGATGATGCCAA 2687
 QY 561 GluValSerAlaIa 580
 DB 2688 GAAGTGTACCTGACACACAGGGCCCGCTACTTCTGGGGTAACTTCCGGATGAACAGG 2747
 QY 581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGluCysLeuGluHisGlyArg 600
 DB 2748 CCGTTGGCATTCACCTGGAATGATTAAGCTGAGCTGAGAGTGTGTGGAGCATGGACG 2807
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 DB 2808 ATACCAAGTTCCACAAAGTGAAGACCATTAACGAGTCAACATCTCAATAAACAGGGC 2867
 QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
 DB 2868 AAGACCAAGCATTTTCTGTTCTTCATGATGAGAAAGAGACATCTTAATGTCACATGAA 2927
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2928 ATGGAAAGGATATTGGTTTCCAGTCCACTATACGATCACTTCCAAATGAGCCGCTTG 2987
 QY 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
 DB 2988 GCGAGGAGAGACTGCTGGGCGGCTATGAGCGTGCATCATCCGCCACTTTCGCT 3047
 QY 681 ProLeuLysGluTyrPheAlaCysVal 689
 DB 3048 CCGCTGAAGAGATATTTCGCTGTGTG 3074
 RESULT 9
 QY CQ720469 2841 bp DNA linear PAT 03-FEB-2004
 DB LOCUS CQ720469
 DEFINITION Sequence 6403 from Patent WO2068579.
 ACCESSION CQ720469
 VERSION CQ720469.1 GI:42281326
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 AUTHORS Kits, such as nucleic acid arrays, comprising a majority of
 TITLE humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 6403 06-SEP-2002;
 FEATUES PE Corporation (NY) (US)
 source Location/Qualifiers
 1. 2841
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2841
 Score: 3746.00 Matches: 689
 Percent Similarity: 98.0% Conservative: 0
 Best Local Similarity: 98.0% Mismatches: 0
 Query Match: 99.5% Indels: 14
 DB: 2 Gaps: 1
 US-10-623-813-86 (1-689) x CQ720469 (1-2841)
 QY 1 MetAsnAlaValGluGluAsnGlnGlyProGlyLysSerGlnLysValGluGluAlaSer 20
 DB 730 ATGAATCTGTGGAAAGAAACCAAGGGCCCGGAGAGTCTCAAGAGGTGAGAGGCCACG 789
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40

Db 790 CCTCCTGCTGTCAGCAGCCACCTGACCCCGCATCCCGCATGTCGCTACCAACGCGCTTAC 849
 Qy 41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
 Db 850 CCCCTGGGGTCCGATGCTGGGGAGCAAGAAATCCACCAAGCAGCGGAGACGAGCAGAG 909
 Qy 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
 Db 910 TACAGAGACGCGCGGGCTTTGGCATTTGGAGCTGTGTGGGGAACTCGGGGGCTTC 969
 Qy 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
 Db 970 TCTCGTGGCCAGCGCCGATTTGTGTCGATGAGCGGCCGAGCCGAGCGAGCTGAA 1029
 Qy 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 Db 1030 GGCAACCGCTGGGTCACTGTGTGGAGACCGCAAAATCTCATGTGTGTGTGAGAG 1089
 Qy 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGluPro 140
 Db 1090 CTGATGCGGCTGAGCTGATGTTTGCAGTGCTTCCACAGGCCAGGTACAAACAGCAGCC 1149
 Qy 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 Db 1150 ATGTACCGCAAGCCATCTACGAGGTCTCTGCAAGTGGCCAGACCGCGGGAGAGCTG 1209
 Qy 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
 Db 1210 TTCCCGGTGGCCAGCAGAGGATGAGATGACATCTCCAAAGCGCTGGAGTGCAGAAC 1269
 Qy 181 LysProMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 Db 1270 AAGCCCATGTGATGATGGCCCTGGGGGCTTCCAGCCCTTGTGCGCTTAAGGGCTGAG 1329
 Qy 201 ProProGluGluGlyAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
 Db 1330 CCACCGAAGAGAGGAATCCCTACAAAGATGTACAGGACATGTGGTGGAACT 1389
 Qy 221 GluAlaAlaAlaTyrAlaProProProAlaLysLysProArgLysSerThrAlaGlu 240
 Db 1390 GAGGACACTCTACGACCACTCCACAGCCAAAGAGCCCGGAGAGACAGCGGAG 1449
 Qy 241 LysProLysValLysGluIleLeuAspGluArgThrArgGluArgLeuValTyrGluVal 260
 Db 1450 AAGCCCAAGGCAAGGAGATATTGATGAGCGCAAGAGAGCGGTGGTATGAGAGTG 1509
 Qy 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 Db 1510 CGCGAGAGAGCGCGAGACATTGAGAGCATCTGCTCTGTGGAGCCCTCAATGTTACC 1569
 Qy 281 LeuGlnHisProLeuPheValGlyLysMetCysGlnAsnLysLysAsnCysPheLeuGlu 300
 Db 1570 CTGGAAACACCCCTCTTCTGTGGAGATGTGCCAAACTGCAAGAACTGCTTCTGAG 1629
 Qy 301 CysAlaTyrGlnLysAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
 Db 1630 TGGCGGTACAGTACAGCAGCAGCGCTACCAAGCTCACTGCACTCTGTGGGGGC 1689
 Qy 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 Db 1690 CGGAGAGTGTCTATGTGGGAAACAAACACTGTCGAGCGTCTTTGCGTGGAGTGTG 1749
 Qy 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTyrPasnCys 360
 Db 1750 GACCTCTTGGTGGGGCGGGGGCTGCCAGGCACTTAAGAGAAAGCCCTGGAATGTC 1809
 Qy 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
 Db 1810 TATATGTCGGGCAAGGGTACTTACGGGCTGTGTGGGGCGCAGAGGACTGGCCCTCC 1869
 Qy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400

Db 1870 CGGCTCAGATGTTCTTGCTTAATAACCAAGCAGCAAGGAAATTTGACCTCCAAAGTTTAC 1929
 Qy 401 ProProValProAlaGlyLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
 Db 1930 CCACCTGTCCAGGTGAGAGAGAGAAACCATCGGGTGTGTCTCTTCTTGAAGAAATC 1989
 Qy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnAlaAspArgTyrIleAlaSer 440
 Db 1990 GCTACAGGCTCTCTGTGCTGAGAGACTTGGGCAATTCAGGTGACCGCTCATTTGCCCTC 2049
 Qy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 Db 2050 GAGGTGTGAGGACTTCATCAGGTGGCATGTGGCGGACCAAGGAGGAGATCATGTAC 2109
 Qy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
 Db 2110 GTGGGGACCTCCGACAGCTCACAGAACATATCCAGAGAGTGGGGCCCATTCGATCTG 2169
 Qy 481 ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 Db 2170 GTGATGGGGGAGTCCCTCATATGACCTTCCATGTCAACCTGTCCAAAGGGCTTC 2229
 Qy 501 Tyr-----GluGlyThrGlyArg 506
 Db 2230 TACGGTACAGCCCGCAGTGTGCTTCTTCCGACCTTCAGAGGGCAGCTGGCCGG 2289
 Qy 507 LeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArg 526
 Db 2290 CTCTTCTTGAAGTTCACCCCTTCGATGATGGCGGGCCAAAGGAGGAGATGATGCG 2349
 Qy 527 ProPhePheTrpLeuPheGluAsnValAlaAlaMetGlyValSerAspLysArgAspIle 546
 Db 2350 CCTTCTTCTGCTCTTTGAGAAATGTGGTCCATGGCCCTTGTGACAGAGGAGATC 2409
 Qy 547 SerArgPheLeuGluSerAsnProValMetIleAspAlaLysGluValSerAlaHis 566
 Db 2410 TCCGATTTCTGAGTCCAAACCTGTGATGATGATGACCAAGAGTGCAGCTGCACAC 2469
 Qy 567 ArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrVal 586
 Db 2470 AAGGCCCGCTACTTCTGGGGTAACTTCCGGTATACAGCGCGGTGGCATCCACTGTG 2529
 Qy 587 AsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyValArgIleAlaLysPheSerLys 606
 Db 2530 AATGATAGCTGAGCTGCAGAGGTCTTGGACATGCGACAGATAGCCAAATTCAGCAA 2589
 Qy 607 ValArgThrIleThrThrArgSerAsnSerIleLysGlnGlyLysAspGlnHisPhePro 626
 Db 2590 GTGAGGACCATTTACTACGAGTCAAACTCCATTAACAGGGCAAGGCCAGCATTTTCT 2649
 Qy 627 ValPheMetAsnGluLysGluAspIleLeuTrpCysThrGluMetGluArgValPheGly 646
 Db 2650 GTCTTCATGATGAGAAAGAGGACATCTTATGTGTGACTGCAAAATGAGAAAGGATTTGT 2709
 Qy 647 PheProValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeu 666
 Db 2710 TTCCAGTCCACTATATCTGAGCTCTCAACATGAGCCGCTTGGCGAGGAGACATGCTG 2769
 Qy 667 GlyArgSerTrpSerValProValIleArgHisLeuPheAlaProLeuLysGlyTyrPhe 686
 Db 2770 GGGCGGTCAATGAGGCTGCAGATCATCCGCACTTGTGCTCGGTGAAGAGATATTTT 2829
 Qy 687 AlaCysVal 689
 Db 2830 GCGTGTGTG 2838
 RESULT 10
 AF480164 2318 bp mRNA linear ROD 14-OCT-2002
 LOCUS AF480164
 DEFINITION Mus musculus DNA cytosine methyltransferase 3A2 (Dnmt3a2) mRNA,
 complete cds.
 ACCESSION AF480164
 VERSION AF480164.1 GI:23954439

| Db | Accession | Source | Organism | Reference Authors |
|------------|--|--|----------|---|
| Db | 1408 | CGTACAGGGCTCTCTGTCTGTCTGAAGACCTGGGACATCCAAAGTGGACCGCTACCTTCCCTCC | | |
| Qy | 441 | GIUVALCYSGIUAASPSETILETHRVAGIUMETVALARGHISGLINGIULYSILEMETYR | | |
| Db | 1468 | GAGGTGTGTGAGGACTCCATCCATCGGTGGGCAATGTGCGGCACCAAGGAAAGATCATGTAC | | |
| Qy | 461 | VALGIYAARVALARGSERVALTHRTGLNYSHISILEGLINGIULTRPGIYPROPHASPLEU | | |
| Db | 1528 | GTCGGGGACGTCCCGACGGCTCACAGAAAGCATATACAGAGATGGGGCCATTCGACCTG | | |
| Qy | 481 | VALIIEGLYGLYSETPROCYASNAAPLEUSERILEVALASNPVALARGULYSIYLEU | | |
| Db | 1588 | GTGATTTGAGGACATCCCTTCGACATATACCTTTCATTTGTCAACCCCTGGCCGCAAGGACATT | | |
| Qy | 501 | TYRGIUGIYTHRGIVARGLEUPHAEHGLUPHETRYRARGLEUENHISAPVALAARPRO | | |
| Db | 1648 | TATAGAGGATACGTGGCGGCTCTTTTGAGTTCTACCGGCTCTGCATGATGCGCGGCC | | |
| Qy | 521 | LYSGIUGIYAARPARARGPROPHETRIEUPHIEGLIUSNAVALVALAMETGLYVAL | | |
| Db | 1708 | AAGGAGGAGATGATCGGCCCTTCTTCGGCTCTTTGGAATGTGGTGGCCATGGCGCTT | | |
| Qy | 541 | SERAPRYVAARGAPRIIESETRAPHETLEUGIUSERANPROVALMEIIEASPAVALYS | | |
| Db | 1768 | AGTGCACAGAGGACATCTCGGATTTCTTGAGTCAACCCGTGATATTGACGCCAA | | |
| Qy | 561 | GIUVALSERIALAALAHARGVALARGTYRPHETRPGIYASMLEUPROGLIYMETASNA | | |
| Db | 1828 | GAAGTGTCTCTGCACACAGGCCCGCTTACTTCTGGGTAACCTTCCTGGCATGAACGG | | |
| Qy | 581 | PROLEUALASERTHVALASNAAPLYSLEUGIULEUGINGIULYSLEUGLUNHISGLYAR | | |
| Db | 1888 | CCTTTGGCATCCACTGTGAATGATTAAGCTGGAGCTGCAAGAGTGTCTGGACACCGGCAGA | | |
| Qy | 601 | IIEALALYSPHESERTLYEVALARGPHRIETHTHRAARGSERASSETILEYSEGLNLY | | |
| Db | 1948 | ATAGCCAAAGTTCACGAAAGTAGGAGCAATACACACAGGTCAAACTCTATTAAGCAGGGC | | |
| Qy | 621 | LYSAPSGINHSHPHETPROVALPHMETASNGIULYSGLIAPRIELEUTRPGYSTRGLU | | |
| Db | 2008 | AAAGCCACGATTTCCCGCTCTTCAAGAAGGAGGAGACATCTGTGGTGCATGAA | | |
| Qy | 641 | METGIUARGVALPHIEGLYPHETPROVALHISYTRTHRAEPVALSERASMETSERARGLEU | | |
| Db | 2068 | ATGGAAGAGGTCTTTGGCTTCCCGCTCCACTACAGACGCTCCCAACATGAGCGCGCTG | | |
| Qy | 661 | ALAAAGGINAAGLEULEUGIYARGSETRIPSERVALPROVALIIEARGHISLEUPHELA | | |
| Db | 2128 | GCGAAGCAGAGATCGCTGGCGCGGATCGTGAGACGTCGGGATATCGGCACCTTTCCT | | |
| Qy | 681 | PROLEULYSGLIYTRPHEALACYSVAL | | |
| Db | 2188 | CCGCTGAAGAAATATTTGCTTTGTGTG | | |
| RESULT 11 | BC007466 | 4094 bp | mRNA | linear |
| LOCUS | BC007466 | | | ROD 29-JUN-2004 |
| DEFINITION | | | | Mus musculus DNA methyltransferase 3a, transcript variant 1, mRNA |
| ACCESSION | BC007466 | | | (GenBank clone MGC:5662 IMAGE:3492853), complete cds. |
| VERSION | BC007466.1 | | | GI:13938620 |
| KEYWORDS | MGC: | | | |
| SOURCE | Mus musculus (house mouse) | | | |
| ORGANISM | Mus musculus | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| | Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; | | | |
| | Sciurognathi; Muridae; Muridae; Murinae; Mus. | | | |
| | 1 (bases 1 to 4094) | | | |
| | Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., | | | |
| | Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., | | | |
| | Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., | | | |
| | Hopkins, R.F., Jordan, H., Moore, T., Wang, S.I., Wang, J., Hsieh, F., | | | |

| TITLE | JOURNAL | PUBMED REFERENCE | AUTHORS |
|----------|---|---|--|
| 1 | Strausberg, R. | Submitted (01-MAY-2001) | National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA |
| 2 | (bases 1 to 4094) | 12477932 | |
| 3 | NIH-MGC Project URL: | http://mgc.nci.nih.gov | |
| 4 | Contact: MGC help desk | Email: cgaps@rmail.nih.gov | |
| 5 | Tissue Procurement: | Gilbert Smith, Ph.D. | CNA Library Preparation: Life Technologies, Inc. |
| 6 | DNA Sequencing by: | Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 | Web site: http://www.shgc.stanford.edu |
| 7 | Contact: | (Dickson, Mark) mcd@paxill.stanford.edu | Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M. |
| FEATURES | SOURCE | clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov | Series: IRAP Plate: 8 Row: n Column: 11 |
| 1 | Location/Qualifiers | This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6681208. | |
| 1. .4094 | /organism="Mus musculus" | | |
| 2 | /mol_type="mRNA" | | |
| 3 | /strain="Czech II" | | |
| 4 | /db_xref="taxon:10090" | | |
| 5 | /clone="MGC:5662 IMAGB:3492853" | | |
| 6 | /tissue_type="mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV." | | |
| 7 | /clone_1ib="NCI CGAP_Lu29" | | |
| 8 | /lab_host="DH10E" | | |
| 9 | /note="vector: pCMV-SPORT6" | | |
| 10 | 1. .4094 | | |
| 11 | /gene="Dnmt3a" | | |
| 12 | /db_xref="GeneID:13435" | | |
| 13 | /db_xref="MGI:1261827" | | |
| 14 | 245. .2971 | | |
| 15 | /gene="Dnmt3a" | | |
| 16 | /codon_start=1 | | |
| 17 | /product="DNA methyltransferase 3A, isoform 1" | | |
| 18 | /protein_id="AAH07466.1" | | |
| 19 | /db_xref="MGI:1393621" | | |
| 20 | /db_xref="GeneID:13435" | | |
| 21 | /db_xref="MGI:1261827" | | |
| 22 | /translation="MPSSGGGDTSSSLEREDDRKGEBOENRKEEROEPSATARK VGPGRKRHPPEVSDTPDPVAVITTKSPPMADDSQSDLPGLNGLEKREPOPEESS PAAGCGAPAEAGEGTETPEPAEAVENGCCVTKEGSGASAGECKEPINISMKME GSGRGRLGIGWESSLRORPMLPTFOAGDPYVISRKREDENTLARMKREAKKAIVYA VMVAVENQASGSQCYVEASPPAAVOQPTPDPAIPYATITPEPVGDAQGNATRYAAD | | |

QY 621 lyeaspglnhisbheprovalphmetasngluysgluaapilieleutrcyethrglu 640
 DB 2762 AAGACACGACATTTCCCGCTTTCATGACGAGGAGGACATCTCGTGTGACTGAA 2821
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2822 ATGAAAGGCGTTTGCTTCCCGCTCCACACACAGCGCTCCACACATGAGCGCTTG 2861
 QY 661 AlaArgGlnArgLeuLeuGlyArgSerTyrSerValProValIleArgHisLeuPheAla 680
 DB 2882 GCGAGCAGACACTGCTGGCGGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2941
 QY 681 ProLeuArgGluTyrPheAlaCysVal 689
 DB 2942 CCGCTGAAGAAATTTTGTCTTGTCG 2968
 RESULT 12
 AF068625 4192 bp mRNA linear ROD 06-DEC-1999
 LOCUS AF068625 Mus musculus DNA cytosine-5 methyltransferase 3A (Dnmt3a) mRNA,
 DEFINITION complete cds.
 ACCESSION AF068625.2 GI:6449467
 VERSION AF068625
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 4192)
 TITLE Okano, M., Xie, S., and Li, E.
 AUTHORS Cloning and characterization of a family of novel mammalian DNA
 JOURNAL (cytosine-5) methyltransferases
 PUBMED Nat. Genet. 19 (3), 219-220 (1998)
 9662389
 2 (bases 1 to 4192)
 REFERENCE Xie, S., Okano, M., and Li, E.
 AUTHORS Direct Submission
 TITLE Submitted (28-MAY-1998) CYRC, Mass. Gen. Hospital, 149 13th Street,
 JOURNAL Charlestown, MA 02129, USA
 3 (bases 1 to 4192)
 REFERENCE Okano, M., Chijiwa, T., Sasaki, H., and Li, E.
 AUTHORS Direct Submission
 TITLE Submitted (04-NOV-1999) CYRC, Mass. Gen. Hospital, 149 13th Street,
 JOURNAL Charlestown, MA 02129, USA
 REMARK Sequence update by submitter
 COMMENT On Nov 18, 1999 this sequence version replaced gi:3327977.
 FEATURES
 source
 1. 4192
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /chromosome="12"
 /map="4.0 cM"
 1. 4192
 /gene="Dnmt3a"
 /gene="Dnmt3a"
 1. 4192
 /note="similar to EST sequences deposited in Genbank"
 Accession Numbers AA052791, AA111043, AA154890, AA240794,
 AA756653, W58898, W59299, W91664, and W91665"
 217. 2943
 /gene="Dnmt3a"
 /function="de novo DNA methylation"
 /note="contains Cys-rich region"
 /codon_start=1
 /product="DNA cytosine-5 methyltransferase 3A"
 /protein_id="AAC0177.2"
 /db_xref="GI:6449468"
 /translacion="MPSGPGDTSSTSLEREDRKEGEDEENKGEKEDEPSATARK
 VGRGRKRKIPVPSSTDPKDPATVTKTSCMAODSGSDLLPNDLKRSEEPDEES
 PAAGKKGAPAEAGGCTETPEASRAVENGCCTVKERGAAGSGKEGKQTNISMKE

GSRGRRLGIGWESSLRORPMPLRFOAGDPVYISKRKDEMLARWKEAEKAKVIA
 VNVAVEENASGSSOKVEASPPAYOOPDPAETVATPEPAGDAGDANATKRAAD
 EPEYERGRGCEIYELVTKLRGSSWMPGRVSMHMTGRSAAGCTRMWPFCKKSEV
 VCEKMLPSSFCSAFHQATNTKQPMYRKAIVLVOVASSRAGLPACHDSBDSG
 KAVEVONKQMIEMALGQFQSPGRGLPEEENKPFKEYVTDMWVEEALVAPPRA
 KKRKSTTEKPKYKEIIDERTRELYVEVQKRNEDICISGSLFVEEALHPLFIG
 MCNCKNPLECAYOYDDGOSYCTICGGRVYLMGNNGNCCRCFCEVDLLVYG
 AACAIKEDPMNVMCGHKGTYGLRRREMPRLQMPFANNDOEPDPKVPVPA
 ERKRPVYSLSPFGIATGLVLKDLGSIQDRIYASEVCDSITVGVYRHQKITYVD
 VRSVTKHIOEWGPFDLVIGSSPCNDLSIVNPKGLYBGTGLPFPFLLHDARK
 EGDRPFEMLFENVVAMGVS DKRDISRFLESNVMTDAKESAAHRAVYWGMLPGNN
 RPLASTVNDKLEIQLERLAKFSKVRTITRNSVSKGKXOHPFVPMNEEDILW
 CTEMERVFGPVIYTDVSNMSRLARQRLIGRSWSVPVIRLLFAPLKEYFACV"

ORIGIN

Alignment Scores:
 Pred. No.: 0
 Score: 3714.00
 Percent Similarity: 98.8%
 Best Local Similarity: 98.5%
 Query Match: 98.7%
 DB: 6
 Matches: 679
 Conservative: 2
 Mismatches: 8
 Indels: 0
 Gaps: 0

US-10-623-813-86 (1-689) x AF068625 (1-4192)

QY 1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnYsValGluGluAlaSer 20
 DB 874 ATAAATGCTGTGGAAGAAACCAAGGCTCTGGAGAGCTAGAAAGGTGAGAGGCCAC 933
 QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
 DB 934 CTTCTGCTGTGACGACGACCCACGAGCCCTGCTTCCGACTGTGGCCACACCCCTGAG 993
 QY 41 ProValGlySerAspAlaGlyAspIlyAsnAlaThrLysAlaGlyAspAspGluProGlu 60
 DB 994 CCAAGTGAAGGCGATCTGGGAGCAAGAAATGCTACCAAGACCCACGATGAGCCCTGAG 1053
 QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuAAspGlyPhe 80
 DB 1054 TATGAGATGCGCGGAGCTTTGGCATTTGGAGACTGGTGGGGGAAACTTCGGGCTTC 1113
 QY 81 SerTyrTrpProGlyArgIleValSerTyrTrpMetThrGlyArgSerArgAlaAlaGlu 100
 DB 1114 TCCTGTGGCCAGACCGCAATTGTGTCTTGGTGATGACAGCGGAGCCGAGCGCTGAA 1173
 QY 101 GlyThrArgTyrValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
 DB 1174 GGCACCTGCTGGGCTCATGTGTCGAGATGGCAAGTTCACGTGTGTGTGTGAGAAAG 1233
 QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
 DB 1234 CTATGCGCTGAGCTCTCTTGTCAAGTGCATTCACACGAGCCACCTACACAGACACCC 1293
 QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
 DB 1294 ATGTACCGCAAGAACCATCTACAGAAATCTTCAGGTGGCCAGCGCTGCGGAACTG 1353
 QY 161 PheProValCysHisAspSerAspGluSerThrAlaLysAlaValGluValGlnAsn 180
 DB 1354 TTTCGAGCTTGCCATACAGTGTGAAGGACAGTGGCAAGCTGTGGAAGTGCAGAAC 1413
 QY 181 LysProMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
 DB 1414 AACCAAGTATTAAGGGGCTCGGAGCTTCCAGCTTCGAGCTCTTAAGGAGCTGAG 1473
 QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTyrValGluPro 220
 DB 1474 CCACCAAGAAAGAAAGAAATCTTCAAGGAAGTTTACCCGACATGTGGTGGAGCCT 1533
 QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysAspArgLysSerThrAlaGlu 240
 DB 1534 GAAGCAGCTGCTTACCCGCCACCCACGACCCCAAGAAACCCAGAAAGACACAAACAGAG 1593

QY 241 LysProValysGluLeileaspGluArgThrArgLysLeuValTyrGluVal 260
 DB 1594 AAACCTAAGGTCAGAGGATCATTCATAGCGCAAGAGCGCGTGTGTATGAGGTG 1653
 QY 261 ArgGlnLysCysArgAsnileGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 DB 1654 CGCAGAGGATGCAGAAACATCGAGGACATTTGATCTCATGTGGGAGCCTCAATGTCAAC 1713
 QY 281 LeuGluHisProLeuPheValGlyClyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
 DB 1714 CTGGAGACCCCACTTCATAGGAGCATGTGCCAGAACTGTAAAGACTGCTTCTGGAG 1773
 QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGly 320
 DB 1774 TGTCTTACCGATGATGACGATGGTGTACCACTTTCATTCACCATTCGCTGTGGGGG 1833
 QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
 DB 1834 CGTGAAGTCTCATGTGTGGGAACAACTGCTGCAAGTCTTTGTGTGAGTGTG 1893
 QY 341 AspLeuLeuValGlyProGlyValAlaGlnAlaIleLysGluAspProTyrAsnGly 360
 DB 1894 GATCTCTTGTGGGGCCAGAGCTGTCTCAGCAGCATTTAAGAAAGACCTCGAATCTGC 1953
 QY 361 TyrMetCysGlyIleHisGlyTyrThrTyrGlyLeuLeuArgArgGlyAspTyrProSer 380
 DB 1954 TACATGTGCGGGCATTAAGGCACTTATGGCTGTCTGCCAAGACGGGAAGACTGGCTTCT 2013
 QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisaspGlnGluPheAspProCysValTyr 400
 DB 2014 CGACTCCAGATGTTCTTTGCCAATACCATGACACGAAATTTGACCCCCCAAGGTTTAC 2073
 QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyLe 420
 DB 2074 CCACCTGTGCCAGCTGAGAAAGAAAGGCCCAATCGCGTGTCTCTTTGATGGGAGT 2133
 QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAsnSer 440
 DB 2134 GCTACAGGGCTCCTGTGTCTTAAGCACTGGGCATCCAAAGTGAACCGCTACATTCCTCC 2193
 QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 DB 2194 GAGGTGTGTGAGCACTCATACCGGTGGCATGTGTGGCGCACAGGAAATCATCTATAC 2253
 QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
 DB 2254 GTCGGGGACGTCCCGACAGGTCAACAGAAAGATATCCAGAGTGGGGCCCATTCACCTG 2313
 QY 481 ValIleGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 DB 2314 GTGATTTGAGGCAAGTCCCTGCATATGACTCTTCATGTTCMAACCTGCGCGCAAGGACTT 2373
 QY 501 TyrGlnGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisaspAlaArgPro 520
 DB 2374 TATAGAGGTACTGACCGGCTCTTCTTGAATGTACCCGCTCTGCATGATGCGGCGCC 2433
 QY 521 LysGlnGlyAspAspArgProPhePheThrLeuPheGluAsnValValAlaMetGlyVal 540
 DB 2434 AAGAGGAGATGATGATCGCCCTTCTTGTGGCTTTTGAAATGTGTGGCCATGGGCTT 2493
 QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
 DB 2494 AGTACAAAGAGGACATCTCCGATTTCTTGAAGTAAACCCCGATGATGATGACGCCAA 2553
 QY 561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
 DB 2554 GAACTGTCTGTGACACAGAGCGGCTTACTTGTGGGATAACCTTCCGTGATGACAGG 2613
 QY 581 ProLeuLysSerThrValAsnAspLysLeuGlnGluCysLeuGlnHisGlyValArg 600
 DB 2614 CTTTGGATTCACCTGTGAATGATGATGCTGGAGGCTGCAAGAGTGTCTGGACACGGCGGA 2673
 QY 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620

DB 2674 ATAGCCAAAGTTACAGCAAGTGAAGGACCATTAACCAAGTCAAACTCTATAAGCAGGGC 2733
 QY 621 LysAspGlnHisPheProValPheMetAsnGlnLysGluAspIleLeuTyrProCysThrGlu 640
 DB 2734 AAAAGCCAGCATTTCCCGCTTCATAGACAGAAAGAGGACATCTGTGTGACCTGAA 2793
 QY 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 DB 2794 ATGGAAGAGGTGTTGGCTTCCCGTCACTACACAGACGCTCTCCAAATAGAGCGGCTG 2853
 QY 661 AlaArgGlnArgLeuLeuGlyArgSerTyrPserValProValIleArgHisLeuPheAla 680
 DB 2854 GCGAGCAGAGACTGTGGCCGATCGTGAAGCGTGCCTGCATCCGCCACTTGGCT 2913
 QY 681 ProLeuLysGlnTyrPheAlaCysVal 689
 DB 2914 CCGCTGAAGGAATATTTGCTGTGTG 2940
 RESULT 13
 LOCUS AY271299 2597 bp mRNA linear MAM 05-SEP-2003
 DEFINITION Bos taurus DNA methyltransferase 3a isoform 4 (DNMT3A) mRNA,
 partial cds.
 ACCESSION AY271299
 VERSION AY271299.1 GI:32350982
 KEYWORDS
 SOURCE Bos taurus (cattle)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 2597)
 AUTHORS Golding,M.C. and Westhusin,M.E.
 TITLE Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and
 Expression in Bovine Preimplantation Embryos, Fetal and Adult
 Tissues
 JOURNAL Gene Expr. Patterns 3 (5), 551-558 (2003)
 REFERENCE 2 (bases 1 to 2597)
 AUTHORS Golding,M.C. and Westhusin,M.E.
 TITLE Direct Submission
 JOURNAL Submitted (08-APR-2003) Veterinary Physiology, Texas A&M - CVM, 500
 University Dr. West, College Station, TX 77843-4466, USA
 FEATURES
 source location/Qualifiers
 1..2597
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 <1..2597
 /gene="DNMT3A"
 <1..2467
 /gene="DNMT3A"
 /note="DNMT3a4; DNA (cytosine 5) methyltransferase"
 /codon_start=2
 /product="DNA methyltransferase 3a isoform 4"
 /protein_id="AAP75902.1"
 /translation="BGEDEEARAKEEROBOSTTARKVGRGRRKXHPVSSDTPDK
 PATTSKSLMAQDSGSEELTPNDLERSPOEBEGSPAGOCAGAPABEGKATETPD
 EASRAVENGCCTREPLOKAEKKAKYIAVMNAVEHQGSTESQKEASBPAAVOQT
 DPASPVTATPPEVPGADAGDKNTKAADPEVEDGFGIIGLWGLKGFSPWPKR
 IVMWMTGSRRAAGETRWMMFGDKFSVVCVSKLMPLSFCSAFHOATYKQPMYRK
 AIVEVLOVASBRAKGLPFWCHDDESDTAAVEVONKOMIEMALGFGPQSGPKLEP
 EEEKNPYKEVYTMVBEPEAAAYAAPPAKPKRSTTEKPKVEIIDERTRRLVVEY
 ROKCRNIEDICTCGSLANTLBPRLFTGNCQCKNGLECAVOYDDGYOSTTIC
 GGRBEVLMCGNNCCRCFCVCDLVLGPGAQAIKEDPMNCTMGKGTGILRRD
 DMSRRLQMPFANNHDOEFDPKXYPVPAKRPKIRVLSLFDGATGLVKDGLGV
 DRYIASVECDSTTVGVNVRHQKIMYGVDSVYTKXIDQSGPDLVIGSPCNDLSI
 VNPARKGLYEGRLEFEEFYRLHDAAPKSGDRPFMWLENNVVLGVSDKSDISRL
 ESNPMTDAKEVDAARBARVFMGNLPLPMNRPLASTVUDKLELOECLEHGRIRAKFSKYR
 TITTRNSITQSGDOHFRPYPMNEKEDILMCTEMERVGVFVHTVDVSNMRLLARQLL
 GRSMSPVILHLPALKEIFYACV"

ORIGIN

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Prod. No.: | 0 | Length: | 2597 |
| Score: | 3712.00 | Matches: | 678 |
| Percent Similarity: | 99.1% | Conservative: | 5 |
| Best Local Similarity: | 98.4% | Mismatches: | 6 |
| Query Match: | 98.6% | Indels: | 0 |
| DB: | 14 | Gaps: | 0 |

US-10-623-813-86 (1-689) x AY271299 (1-2597)

```

Oy      1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGluValGluGluAlaSer 20
Db      398 ATAAATGCTCTTAAAGAAACCAAGGCTCCACCGAGCTCAGAAAGTGGAGAGCCACT 457

Oy      21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      458 CTTCCCGCTGTGCAGAGCCACCGACCTGCGTCCCAACAGTGGCCACCGCCTGAG 517

Oy      41 ProValGlySerAspAlaGlyAspAlaThrThrValaGlyAspAspGluProGlu 60
Db      518 CCGGTGGGGCCGATGCCGGGGAAGAAATGCCAAGACAGCTACGATGAACCGGAG 577

Oy      61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgLysPhe 80
Db      578 TACGAGGAGCGCGGGCTTTGGCATTTGGGAGCTGGTGGGGAAACTGGCGGCTTC 637

Oy      81 SerTyrTrpProGlyValArgIleValSerTyrTrpMetThrGlyValrGSerArgAlaIleGlu 100
Db      638 TCTGTGTGGCCAGCGCCGATTTGTCTGTGTGATGACGGGCGGAGCCGAGCCGAG 697

Oy      101 GlyThrArgTyrValMetTyrPheGlyLysPheSerValValCysValGluLys 120
Db      698 GGCACCTGTTGGGTCAATGTGTTCGAGACGCGCAAGTTCAGTGTGTGTGGAGAG 757

Oy      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db      758 CTGATGCCCTGAGCTCTCTTCAGTGCCTTCCACAGGCGCACTTACAAACAGCAGCCC 817

Oy      141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      818 ATGTACCGCAGAGGCTATCTACGAGGTCTCTCAGGTGCCAGCACTCAGCGGGGAAGCTG 877

Oy      161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluAlaAsn 180
Db      878 TTTCCAAATGTCATGACGACGACGAGAGCGACACTGCCAAGGCCGTGGAGTGCAGAAC 937

Oy      181 LysProMetIleGluTyrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      938 AAACAAATGATCCAGTGGGCCCTGGAGGGTTCCAGCCCTCTGGCCCCCAAGGGCTGGAG 997

Oy      201 ProProGluGluGluLysAspProTyrLysGluValTyrThrAspMetTyrValGluPro 220
Db      998 CCCCAGAGAGAGAGAAAGAACCCCTTACAAAGATTACACAGCATGTGGTGAACCC 1057

Oy      221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db      1058 GAGGAGCTGCTCATGTGGCGGCCCCCAAGCCAAAAGCCCGAAAGACACAAACGAG 1117

Oy      241 LysProLysValLysGluIleIleAspGluArgThrArgLysLeuValTyrGluVal 260
Db      1118 AACCTTAAGGTCAAGAGATCATTTGATGAACGCAAGAGAGCGCGTGTGTACAGGTA 1177

Oy      261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1178 CGGCAGAAAGTGGCGGAACATCGAGACATCTGCTCTTGTGGAGCCCTCAACGTCAAC 1237

Oy      281 LeuGluHisProLeuPheValGlyGlyMetCysGluAsnGlyLysAsnCysPheLeuGlu 300
Db      1238 TTGGAAACACCTCTTCACTGAGGAGATGTGCCAAATGCAAGAACTCTCTTCGAA 1297

Oy      301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320

```

```

Db      1298 TGGCGGTACAGTACGATGAGCTATCAGTCTCACTACGACCACTCTGCGGGGCG 1357

Oy      321 ArgGluValLeuMetCysGlyLysAsnAsnGlyCysArgCysPheCysValGluCysVal 340
Db      1358 CGGAGAGTCTCATGTGTGGAAACAATGTCTGAGGGTCTTTTGGCGGGAAATGTGTG 1417

Oy      341 AspLeuLeuValGlyProGlyAlaIleGlnAlaIleIleGlyLysAspProTyrAsnGly 360
Db      1418 GATCTTGTGGGGCCAGGGGCCGCGAGCGAGCCATCAAAAGAACCCCTGGAAGTGC 1477

Oy      361 TyrMetCysGlyLysLysGlyThrTyrGlyLeuLeuArgArgLysAspTyrProSer 380
Db      1478 TACATGTGGCGGACAAAGGCACTTACGGGCTCTCGGGCGGGAGCGACTGGCCGTCT 1537

Oy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProLysValTyr 400
Db      1538 CCGCTCCAGATGTTCTTCCCAACAACATGACAGAGAAATTCATCTCCGAAGTTTAC 1597

Oy      401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      1598 CCACTGTCCCACTAGAAAGAAAGCCCATCCGGGTGCTCTCTATTGATGGAATT 1657

Oy      421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      1658 GCTACAGGGCTTCTGTGCTGAAGAGACTTGGGCATTACGGTGAACCGCTACATCGCCTCC 1717

Oy      441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      1718 GAGGTGTGGAGACATCCATCAGGTGGGATGTGTGGGACCAAGGAAAGATCATGTAC 1777

Oy      461 ValGlyAspValArgSerValThrGlnLysHisIleGlnIleTyrGlyProPheAspLeu 480
Db      1778 GTGGGGAGCTCCGACGCTTACACAGAACATATCCAGAAATGGGGCCGCTTGATCTG 1837

Oy      481 ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaIleGlyLysLeu 500
Db      1838 GTGATTTGGGGGAGTCCCTGCAATGATCTCTCATGTCTCAACCTCGCCCAAGGGCTC 1897

Oy      501 TyrGluGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      1898 TACGAGGACATGGCGGCTCTTGTAGTTTACCGCTCTGCAATGATGGCGGCC 1957

Oy      521 LysGluGlyAspAspArgProPhePheTyrLeuPheGluAsnValAlaIleMetGlyVal 540
Db      1958 AAGGAGGAGATACCGGCCCTCTCTGCTGTGAGAAATGTGGTGTGCTCGGGGCTT 2017

Oy      541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      2018 AGTGACAAAGGGGACATCTCGCATTTCTCGAGTCCAAACCTGTGATGATGATGCCAAA 2077

Oy      561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
Db      2078 GAAGTGTACGTGGGACAGGGGCCCTCTACTTCTGGGGGAACCTTCTGGTATGAACAGG 2137

Oy      581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnLysCysLeuGluHisGlyArg 600
Db      2138 CCAITGGCATCCACTGTGAATGAATAGCTGAGCTCAGAGAGTGTCTGAGAGACGGCGA 2197

Oy      601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2198 ATAGCCAAAGTTCCAGAAAGTGAAGACCATTACTACAGGTGCAAGTCCATAAAGCGGGC 2257

Oy      621 LysAspGlnHisPheProValIlePheMetAsnGluLysGluValLeuTyrPysThrGlu 640
Db      2258 AAAGACACAGATTTCCCGTCTTCATGAATGAAGAAAGAGACATCTTATGTGCACTGAA 2317

Oy      641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2318 ATGGAAGGGTGTGTTGGCTTCCCTGTCCAATAACGAGTCTCCAACATGAGACCGGTTG 2377

Oy      661 AlaArgGlnArgLeuGluArgSerTyrSerValProValIleArgHisLeuPheAla 680
Db      2378 GCGAGGCAGAGACTGTGGGCGGTGTGGAGCGTCCGGTCAATCGGCCACTTTCGCT 2437

```

```

QY      681  ProLeuLysGluTYrPheAlaCysVal 689
Db      2438 CCGCTGAAGAAATATTGCTGTGCTG 2464

RESULT 14
LOCUS      AY271298      2798 bp      mRNA      linear      MAM 05-SEP-2003
DEFINITION Bos taurus DNA methyltransferase 3a (DNMT3A) mRNA, partial cds.
ACCESSION  AY271298
VERSION     AY271298.1
KEYWORDS    GI:32350980

SOURCE
ORGANISM    Bos taurus (cattle)
            Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
            1 (bases 1 to 2798)
REFERENCE   Golding, M.C. and Westhusin, M.E.
AUTHORS     Golding, M.C. and Westhusin, M.E.
TITLE       Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and
            Expression in Bovine Preimplantation Embryos, Fetal and Adult
            Tissues
            Gene Expr. Patterns 3 (5), 551-558 (2003)
            2 (bases 1 to 2798)
            Golding, M.C. and Westhusin, M.E.
            Direct Submission
            Submitted (08-APR-2003) Veterinary Physiology, Texas A&M - CVM, 500
            University Dr. West, College Station, TX 77843-4466, USA

FEATURES
source
            1..2798
            /organism="Bos taurus"
            /mol_type="mRNA"
            /db_xref="taxon:9913"
            <1..2798
            /gene="DNMT3A"
            <1..2668
            /gene="DNMT3A"
            /note="DNMT3A; DNA (cytosine 5) methyltransferase"
            /codon_start=2
            /product="DNA methyltransferase 3a"
            /protein_id="AAP75901.1"
            /db_xref="GI:32350981"
            /translacion="EGEROERARAKEEROBETTARKVGRGRKKHPVSSDPPK
            PAVTSKLSMAQDSGSELTYPNGDLEKSESEQPEPESGAGCKGAPRAGGATETPR
            EASAVENGCCTPKDGRGAPAEBEKEQETNIESMKESGRLRGGLGWESLQR
            MPRLTFOAGDPRYISKRKDEMLARKWEAKAKAVALVMAVENQSTSQVEEA
            SPAPVOPTDASPVTATTPPVGADAGDKNAATKADADEPEYEDRGFGIGELVWGX
            RGFEWMPGRIVSMWMTGRSRAEGRWMMGDKGSVCVEKMLPSSFCSAFHQAT
            YNKQPMYRKAIYEVLQVASSRAGKLPFKCHSDSDPAKAYENQNKMIEMALGROP
            SGPRGLEPPEEKRPYKEVYTDWVEPAAAIYAPPPAKRKSTTEKPKYKEITDER
            TRELVYEVROKCRNIEDI CISCGLNVTLLHPLEFGMCQCNKCFLECAVOYDDG
            YOSVCTICCGGREVLKMGNNCCRCFCVCEVDLVGPAAQAIKEDPWNVYCHKG
            TYGLLRDDWPSRLQMFANNHDOEPKPYRPPVPEKKAPIRVLSLFGIATGL
            VLKQIGOVDPYVIASEVCEDSITVGMVHOKIMVGVRSVTOGHIOEMFPDLVIG
            GSPNDLSIVNPAKGLYEYCGRLPFEEYRLLHDAPRPEGDRPFMLFENVVVLGS
            DKRISFRLESNPMTIDAKESAAHRAATYFNGNLPGMNRPPLASTVNDLLEIQECLHS
            RIAPFSKRVITTTNSNISKQGDHPFVFNMEKEDIIMCTEMERVFGEVPHYTVSNM
            SRLRORLLGRHSWSPVIRHLFAPLKEYFACV"

ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      2798
Score:          3712.00      Matches:      678
Percent Similarity: 99.1%      Conservative: 5
Best Local Similarity: 98.4%      Mismatches: 6
Query Match:    14      Indels:      0
                  Gaps:      0

US-10-623-813-86 (1-689) x AY271298 (1-2798)
QY      1  MecAsnaLaValaGluGluAsnGlnGlyProGluGluSerGlnLysValaGluGluAlaSer 20
Db      599 ATGATGCTGTGTAAGAAACCAAGGGGTCCACCGAGTCTCAAGAGCTGAGAGAGGCCAGT 658

```

```

QY      21  ProProAlaValaGlnGlnProThraSPProAlaSerProThraValaAlaThrThProGlu 40
Db      659 CCTCCCTGTCGACGACCCACCGACCTGCGTCCCCACAGAGGGCACACCCCTGAG 718

41  ProValaGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
Db      719 CCCGTGGGCGCCGATGCCGGGGACAAAGATGCCAACAGCTGACGATGAACCCGAG 778

61  TyrGluAspGlyArgGlyPheGlyLeuValaIrrGlyLysLeuArgGlyPhe 80
Db      779 TACGAGACGGCCGGGCTTTGGCATTTGGGAGCTGTGTGGGGAAATCGCGGGCTTC 838

81  SerTrpTrpProGluYargGlyValaSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db      839 TCCGTGGGCGCAGGCGCCATTTGTCTTGTGTGAATGACGGGCGGACCGACGCGAA 898

101  GlyThraArgTrpValMetTrpPheGlyAspGlyLysPheSerValaCysValaGluLys 120
Db      899 GGCACCTGTTGGGTCATGTGTGTCGAGACGGCAAGTTCTCAGTGTGTGTGTGAGAG 958

121  LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyraSnlYsGlnPro 140
Db      959 CTGATGCCGCTGAGCTCTTCTGCAGTGCTTCCACAGGCGCACCTCAACAGCAGCC 1018

141  MetTyraArgLysAlaLeuTyrgluValLeuGlnValaLaseSerArgAlaGlyLysLeu 160
Db      1019 ATGATCCGACAGGCTATCTACAGAGTCTTCAGAGTGCAGCACTCAAGCGGGAGCTG 1078

161  PheProValCysHisAspSerAspGluSerAspThraLysAlaValaGluAlaAsn 180
Db      1079 TTTTCAATGTGCCATGACAGAGACGAGACGACCTCCAAAGCGCGTGAAGTGCAGAC 1138

181  LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      1139 AAACAAATGATCGAGTGGCGCTCGGAGGAGTTCAGCCCTCTGCGCCCAAGGCGCTGAG 1198

201  ProProGluGluGluLysAsnProTrpLysGluValaTyrrThraSPMetTrpValaGluPro 220
Db      1199 CCCCACAAAGAGGAGAAACCCCTCAAGAAAGTTTACACAGACATGTGGTTGAACC 1258

221  GluAlaAlaAlaTyraLysProProProProAlaLysLysProArgLysSerThraAlaGlu 240
Db      1259 GAGGCACTGCTTATGCGCCGCCGCCACACAGCCAAAGCCGAAAGAGCACAACGAG 1318

241  LysProLysValaLysGluIleIleAspGluArgThraArgLysLeuValaTyrgluVala 260
Db      1319 AAGCTAAGGTCAAAGAGATCATTTGATGAACGACACAAGAGCGCGTGTGTACAGGTA 1378

261  ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerIleuAsnValThr 280
Db      1379 CGCGAAGAGTCCGAAATCTGAGAGACATCTGCTCTTGTGGAGCTTCAAGCTCAC 1438

281  LeuGluHisProLeuPheValaGlyLysCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db      1439 TTGGAACACCTCTCTTCACTGAGAGAAATGTCACAAAACCTGCAAGAACTGCTTCTG 1498

301  CysAlaTyrglnTyraAspAspAspGlyTyrglnSerTyrcysThrIleCysGlyGly 320
Db      1499 TGCCGTCACAGTACGATGAGTACGATGCTATACCTCACTGCAACATCTGCTGGGGGG 1558

321  ArgGluValaLeuMetCysGlyLysAsnAsnCysCysArgCysPheCysValaGluCysVala 340
Db      1559 CGCAGGTGCTCATGTGTGGAAACAAATTTGCTGACGGTCTTTTTCGTGGAATGTGTG 1618

341  AspLeuLeuValaGlyProGluValaAlaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db      1619 GATCTCTGTGTGGGCGCAGGGGCGCGCAGGACCATCAAGAAGACCCCTGGAACGTC 1678

361  TyrMetCysGlyHisLysGlyThrTyrglyLeuLeuArgArgArgGluAspTrpProSer 380
Db      1679 TACATGTGCGGCGCAAGGGGACACTGAGGCTGCTGCGGGCGGCGGACGACTGGCCGCT 1738

```

```

Oy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAsnGlnGluPheAspProPheValTyr 400
Db 1739 CGGCTCCAGAGTTCCTCGCCAAACACATGACAGGAATTCCTCCGAAAGCTTAC 1798
Oy 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 1799 CCACCTCTCCAGCTGAGAGAGAAACCCATCCGGGTGCTGCTCTAATTCGATGGAATT 1858
Oy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 1859 GCTACAGAGGCTTCTGCTGCTGAAGAGACTTGCGCATTCAGGTGAGCCCTCATCGCTCC 1918
Oy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 1919 GAGGTGTGTAGAGACTCATTCACGTGGGACATGGTGGGACCCGGAAGATCATATAC 1978
Oy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGlnIleTyrGlyProPheAspLeu 480
Db 1979 GTGGGGGACCTCCGACAGCTTACACAGACATATCCAGGAATGGGCGCGTTCGATCTG 2038
Oy 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db 2039 GTGATTGGGGGAGTCCCTCAATGATCTCTCCATCGTCAACCTCGCCCAAGGAGCATC 2098
Oy 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 2099 TAGGAGGGCAGCTGGCGGCTCTTCTTGAGTTCACCGCTCCGACGATGATGCCGGCCC 2158
Oy 521 LysGluGlyAspAspArgProPhePheTyrLeuPheGluAsnValValAlaMetGlyVal 540
Db 2159 AAGGAGGAGATATACCGCCCTTCTTCTGCTCTTTAGATGATGCTGCTGGGCTT 2218
Oy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 2219 AGTACAGAGGAGGACATCTCGCATTTCTCGAGTCCAAACCTGTGATGATGATGCCAA 2278
Oy 561 GluValSerIleAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
Db 2279 GAAGTGTACAGCTCGACAGAGCCCGCTACTCTCGGGGAACTTCTGCTATGAAACAGG 2338
Oy 581 ProLeuAlaSerThrValAsnAspLysLeuGluGlnGluCysLeuGluHisGlyArg 600
Db 2339 CCATTGGCATCCACTGGAATGATTAAGCTGAGAGTGTCTGAGACACGGCCGA 2398
Oy 601 IleAlaLysPheSerLysValArgThrIleThrThaArgSerAsnSerIleLysGlnGly 620
Db 2399 ATAGCCAAAGTTACAGAAAGTGAAGACCATTACTACTAGTCAACTCATTAAGCAGGGC 2458
Oy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
Db 2459 AAGACCCAGATTTCCCGCTTCATGAAGAAGAAAGAGACATCTTATGCTGACTGAA 2518
Oy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2519 ATGGAAGAGGTGTTTGCTTCCCTGCTCACTATACCACTCTCCAAACAGACCGCTTG 2578
Oy 661 AlaArgGlnArgLeuLeuGlyArgSerTyrPheSerValProValIleArgHisLeuPheAla 680
Db 2579 GCGAGGACAGAGACTGCTGGGCGGCTGTGAGCGGTGCATCCGCCACTTCTGCT 2638
Oy 681 ProLeuLysGluTyrPheAlaCysVal 689
Db 2639 CCGCTGAAGGAATATTTGCTGTGTG 2665

```

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
1 Homiidae; Homo.
TITLE Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A., Yokoyama,S.,
JOURNAL Ohara,O., Nagase,T. and F.Kikuno,R.
AUTHORS None
REFERENCE Published Only in Database (2005)
AUTHORS 2 (bases 1 to 4476)
TITLE Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A., Yokoyama,S.,
JOURNAL Ohara,O., Nagase,T. and F.Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2005) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research, 2-6-7 Kazusa-kametani, Kisarazu,
Chiba, 292-0818, Japan (E-mail:cdnainfo@kazusa.or.jp),
URL:http://protein.gsc.riken.go.jp/, Tel:81-438-52-3930,
Fax:81-438-52-3931)
COMMENT This work was supported in part by the National Project on Protein
Structural and Functional Analysis, Ministry of Education, Culture
, Sports, Science and Technology of Japan. Totoki Y, Toyoda A,
Takeda T, Sakaki Y, Tanaka A, Yokoyama S. RIKEN Genomic Sciences
Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.
e-mail: aketaka@postman.riken.go.jp.
URL: http://protein.gsc.riken.go.jp/.
FEATURES
source
location/Qualifiers
1..4476
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pJ01914"
/tissue_type="brain"
/note="This clone is also named as hsk003000130
vector:pBluescriptII SK plus"
<1..4476
/gene="DNA cytosine methyltransferase 3 alpha isoform a
variant"
<148..2583
/gene="DNA cytosine methyltransferase 3 alpha isoform a
variant"
/inference="non-experimental evidence, no additional
details recorded"
/note="Start codon is not identified."
/codon_start=1
/product="DNA cytosine methyltransferase 3 alpha isoform a
variant"
/protein_id="BAD92070.1"
/db_xref="GI:62087246"
/translation="LSPPKTTIIPSPKSPSPAPSPSPSPGQAMPSPSGPDTSS
SAAEREDKDEBEDEPRKSEBROPSRTTARVVGPGRRKRPVPSGDTPPDAVI
SKSPSAQDSGASBELPFGDLERKSEPOPEBSPAGGOGGAAEBEGAAETLPEAR
AVENGCTPKEBEGAPAEAGKEQETNISMKKEGSRGLRGGLGESSLRQPMPL
TFOAGDPYVTSKRKDEMLARWKREAEKAKYIAGNNVAEENGPEBESQVEBASPA
VOOPTPASPTVAITPEPEVSGDAGDNATKAGDEPEYDGRFGIGELVWGILRGS
WMPGRIVSMQDSGSRRAAGETRVWVWFQGXSVCEVKLMPISFCSAFAHQATVNG
PWYRAIYRVLQVSRARAKLPVCHDSDESPATKAVYONKRMIMMAGGFPQSGPK
GLBPBEKQPIKEVITDWWVEBPAAYIPPPAKPRKSTAEKPYKEIIDRTER
LVYEVAQKRCNIEDICISGSLNVLTLEHLPVFGMKQNKNCLECAVQIDDGYSY
CTICCGREARNICNNCCRCVECDVLLVGGAAQAAIKEDPMNVCYMGHGTYGL
LRRBEDMPRLQWPFANNDQEDFPKYPVPVPAEGRKPIRVSLFDGIATGLVLVKD
LGIQVDVRYASEVCEDSITVGVHRHQKILMYGVDSVSTQKHIOENGPFLLVGGSC
NDLSIYNPARKULYECTGRLLPFEPFYLLHDARPKEDDPRFFLFPFNVAVMGVSDDKD
ISRFLEVNQSHLY"

```

ORIGIN

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Pred. No.: | 0 | Length: | 4476 |
| Score: | 3702.00 | Matches: | 689 |
| Percent Similarity: | 91.1% | Conservative: | 0 |
| Best Local Similarity: | 91.1% | Mismatches: | 0 |
| Query Match: | 98.4% | Indels: | 67 |
| DB: | 5 | Gaps: | 1 |

US-10-623-813-86 (1-689) x AB208833 (1-4476)

QY 1 MetAsna1aValGluGluasnGlnGlyProGlyGluSerGlnLysValGluGlu1aSer 20

Db 907 ATGAATGCTGTGAGAGAAAACAGGGGCCCCGGGAGTCTCAGAAAGTGGAGAGGCCAGC 966

QY 21 ProProValaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40

Db 967 CCTCTGTGTGTGACAGCCCACTAAGCCCGCATCCCTCCAGTGTGATCCACGCGCTGAG 1026

QY 41 ProValGlySerAspAlaGlyAspLysAsna1aThrLysAlaGlyAspAspGluProGlu 60

Db 1027 CCGGCGGGTCCGAGTGTGTGGGACAAAGAAATGCCCAACAGCGGAGTACAGCAGCAGAG 1086

QY 61 TyrGluAspGlyArgGlyPheGlyTlleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80

Db 1087 TACAGAGACGGCGGGGCTTTGGCATTTGGGAGAGTGTGTGGGAGAACTCGCGGGGCTTC 1146

QY 81 SerTrpTrpProGlyArgGlyLeuValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100

Db 1147 TCCTGTGTGCGCAGCGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1206

QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120

Db 1207 GGCACCCCTGGGTATGT 1266

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrThrAsnLysGlnPro 140

Db 1267 CTGATGCCGCTGAGCTGCTTTTGCAAGTGGCTTCCACAGCCACGCTACACACAGAGCC 1326

QY 141 MetTrpArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160

Db 1327 ATGTACCCCAAGCATCTACAGAGTCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1386

QY 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180

Db 1387 TTCCTGT 1446

QY 181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200

Db 1447 AAGCCCATGATTGAATGGGCTCGGGGGGCTTCCAGCTTCTGTGAGCCCTGAGGCTGAG 1506

QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220

Db 1507 CCACCAAGAGAGAGAGAAATCTCTACAAAGAGTGTACACGAGCATGTGGTGGAACTT 1566

QY 221 GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240

Db 1567 GAGGACGCTGCTTACGACACACCTCCACACGCAAAAAGCCCCGGAGAGGACAGCGGAG 1626

QY 241 LysProLysValLysGluLleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260

Db 1627 AAGCCCAAGGTCAAGAGATTATTGTAGAGCCACAGAGAGCGGCTGTGTACACAGGAG 1686

QY 261 ArgGluLysCysArgAsnIleGluAspTlleCysIleSerCysGlySerLeuAsnValThr 280

Db 1687 CGCGCAAGGTGCGGAAACATTGAGACATCTGCATCTCTGTGGAGCCTCATGTGTAC 1746

QY 281 LeuGluHisProLeuPheValGlyLysMetCysGlnAsnCysLysAsnCysPheLeuGlu 300

Db 1747 CTGGAACACCCCTCTTCTGTGTGAGGAATGTCCAAAACCTGCACAAAGCTTTTCTGGAG 1806

QY 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrTlleCysCysGlyGly 320

Db 1807 TGTGTGTACAGTACAGACACAGCGCTACAGTCTTACTGTACACATCTGCTGTGGGGC 1866

QY 321 ArgGluValLleuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340

Db 1867 CGT 1926

QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTyrAsnCys 360

Db 1927 GACCTTGT 1986

QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380

Db 1987 TACATGTGTGGGCAAAAGGATCTTACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2046

QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400

Db 2047 CGGCTCAGAGTCTTCTGT 2106

QY 401 ProProValProAlaGluLysArgLysProLleArgValLeuSerLeuPheAspGlyTlle 420

Db 2107 CCACCTGTCCACGCTGAGAGAGAGAGAGCCCATCCGGGCTGTGTGTGTGTGTGTGTGTGT 2166

QY 421 AlaThrGlyLeuLeuValLleuLysAspLeuGlyTlleGlnValAspArgTyrTlleAser 440

Db 2167 GCTACAGGCTCTGT 2226

QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460

Db 2227 GAGGT 2286

QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrProPheAspLeu 480

Db 2287 GTCCGGGACGTCCCGCAGGCTCACACAGAGCATATCCAGAGTGGGGCCATTGATGTG 2346

QY 481 ValIleGlySerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500

Db 2347 GTGATGTGGGGACATCTCTGCAATGACCTTCTTCATCTGTCAACCTGCTCGCAGAGGCTTC 2406

QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheThrArgLeuLeuHisAspAlaArgPro 520

Db 2407 TACAGGCACTGGCCGCTCTTTGTGAGTTTACCGCTCTGTACATGTGGGGGCC 2466

QY 521 LysGluGlyAspAspArgProPhePheThrLeuPheGluAsnValValAlaMetGlyVal 540

Db 2467 AAGAGGGAGATGATGTGCCCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2526

QY 541 SerAspLysArgAspIleSerArgPheLeu----- 550

Db 2527 AGTGCAGAGAGAGCATCTGCGATTCTTCTCGAGGTACAAAATAGCCACTGTACTAGTGG 2586

QY 550 ----- 550

Db 2587 GCTAGTGGCACTGAATTGAACAGTGTAGAGACAACCCACATCGCAGAGTTCATTTC 2646

QY 550 ----- 550

Db 2647 GACAATGCTGATCTTGAACGTTCCAGTGCAGCTTGTGAAATGTGGCTTAGGTGAATTCC 2706

QY 551 -----GluSerAsnPr 554

Db 2707 AGCTTCTCTGAGGCTGACAGGCTAACCCAGTGTGTGCTCTAGAGAGAAATCCAAACC 2766

QY 554 OvaIleMetIleAspAlaLysGluValIleSerAlaAlaHisArgLlaArgTyrPheTrpGlyAs 574

Db 2767 TGTGATGATTATGCAAAAGAGTGTGCTGCACACAGGGCCGCTACTTCTGGGGTAA 2826

QY 574 nLeuProGlyMetAsnArgProLeuAlaSerThrValAsnAspLysLeuGluGlnGly 594

Db 2827 CTTTCCCGGTATGAACAGGCGTGTGCACTCACTGTGATATATAAGCTGTGAGAGCTGCAGA 2886

QY 594 nCysLeuGluHisGlyArgIleAlaLysPheSerLysValArgThrIleThrThrArgSe 614

Db 2887 GTGTGTGAGCATGTGAGAGATAGCCAACTTACAGAAATGAGACATTACTACAGAGTCT 2946

QY 614 nAsnSerIleLysGlnGlyLysAspGlnHisPheProValPheMetAsnGluLysGluAs 634

Db 2947 AAATCTCAATAAAGCAGGCAAAAGACACAGCATTTCTCTGTTCATGAATGAGAAAGAGA 3006

QY 634 nIleLeuTrpCysThrGluMetGluArgValPheGlyPheProValHisTyrThrAspVal 654

Db 3007 CATCTTATGTGTGCACTGAATGAGAAAGGATATTGTGTGTCCAGTCCACTATACAGAGT 3066

OY 654 1SerPseMetSerArgLeuAlaArgGlnArgLeuGlnArgSerTrpSerValProVa 674
DB 3067 CTCACACATGAGCCGCTGGCGAGGCAAGAGACTGCGCGGTATGAGAGCGTCCAGT 3126
OY 674 11LeArgHisLeuPheAlaProLeuLysGlnTyrPheAlaCysVal 689
DB 3127 CATCGGCCACCTCTTCGCTCCGCTGAGGAGATTTTGGGTGTGTG 3172

Search completed: October 5, 2006, 21:09:00
Job time : 10579 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 5, 2006, 21:00:34 ; Search time 2088 Seconds
(without alignments)
6082.026 Million cell updates/sec

Title: US-10-623-813-86

Perfect score: 3763
Sequence: 1 MNVENVGSGESQKVEAS.....MSVPVIRHLPAPLKEYFACV 669

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 18892170 seqs, 6143817638 residues
Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xjh
-O=/abs/ABSSMBR.spool/US10623813/rnuc 05102006 121153 27814/app_query.fasta_1
-DB=Published Applcations NA Main -OEMT=fastcap -SUFFIX=p2n.rnphm
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=biosum62 -TRANS=human40.cdt -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MNLN=0 -MAXLEN=200000000 -HOST=abs02h
-USER=US10623813 @CGN 1.1 1675 @runat 05102006 121153 27814 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications_NA_Main:
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------------------------|
| 1 | 3763 | 100.0 | 2371 | 9 | US-10-623-813-84 Sequence 84, Appl |

| | | | | | |
|----|--------|-------|--------|----|--|
| 2 | 3763 | 100.0 | 3005 | 6 | US-10-144-577-2 Sequence 2, Appl1 |
| 3 | 3763 | 100.0 | 3005 | 7 | US-10-172-118-132 Sequence 132, App |
| 4 | 3763 | 100.0 | 3005 | 8 | US-10-342-887-132 Sequence 132, App |
| 5 | 3763 | 100.0 | 3009 | 11 | US-10-330-773-279 Sequence 279, App |
| 6 | 3763 | 100.0 | 4293 | 5 | US-09-720-086-3 Sequence 3, Appl1 |
| 7 | 3763 | 100.0 | 4293 | 9 | US-10-623-813-3 Sequence 3, Appl1 |
| 8 | 3714 | 98.7 | 2318 | 9 | US-10-623-813-83 Sequence 83, Appl |
| 9 | 3714 | 98.7 | 4192 | 5 | US-09-720-086-1 Sequence 1, Appl1 |
| 10 | 3714 | 98.7 | 4192 | 9 | US-10-623-813-1 Sequence 1, Appl1 |
| 11 | 3714 | 98.7 | 6157 | 11 | US-10-330-773-276 Sequence 276, App |
| 12 | 2713 | 72.1 | 2938 | 7 | US-10-264-237-953 Sequence 933, App |
| 13 | 2169.5 | 57.7 | 4195 | 5 | US-09-720-086-2 Sequence 2, Appl1 |
| 14 | 2169.5 | 57.7 | 4195 | 9 | US-10-623-813-2 Sequence 2, Appl1 |
| 15 | 2157 | 57.3 | 4145 | 5 | US-09-720-086-4 Sequence 3, Appl1 |
| 16 | 2157 | 57.3 | 4145 | 6 | US-10-144-577-3 Sequence 38, Appl |
| 17 | 2157 | 57.3 | 4145 | 6 | US-10-144-577-38 Sequence 1337, Ap |
| 18 | 2157 | 57.3 | 4145 | 7 | US-10-172-118-1337 Sequence 1337, Ap |
| 19 | 2157 | 57.3 | 4145 | 8 | US-10-342-887-1337 Sequence 4, Appl1 |
| 20 | 2157 | 57.3 | 4145 | 9 | US-10-623-813-4 Sequence 39, Appl |
| 21 | 1878.5 | 49.9 | 3897 | 6 | US-10-144-577-39 Sequence 275, App |
| 22 | 1793.5 | 47.7 | 95484 | 11 | US-10-330-773-275 Sequence 278, App |
| 23 | 1771.5 | 47.1 | 127722 | 11 | US-10-330-773-278 Sequence 926, App |
| 24 | 1318 | 35.0 | 1060 | 6 | US-10-106-698-926 Sequence 1235, Ap |
| 25 | 1095.5 | 29.1 | 709 | 3 | US-09-969-034-1235 Sequence 21969, A |
| 26 | 1003 | 26.7 | 711 | 10 | US-10-450-763-21969 Sequence 272, App |
| 27 | 949 | 25.2 | 632 | 3 | US-09-764-891-272 Sequence 60665, Sequence 122074, |
| 28 | 942 | 25.0 | 994 | 12 | US-10-301-480-60665 Sequence 122074, |
| 29 | 942 | 25.0 | 994 | 12 | US-10-301-480-122074 Sequence 40, Appl |
| 30 | 872.5 | 23.2 | 2057 | 6 | US-10-144-577-40 Sequence 41, Appl |
| 31 | 785 | 20.9 | 2008 | 6 | US-10-144-577-41 Sequence 54, Appl |
| 32 | 750 | 19.9 | 440 | 5 | US-09-720-086-54 Sequence 54, Appl |
| 33 | 750 | 19.9 | 440 | 9 | US-10-623-813-54 Sequence 21970, A |
| 34 | 693 | 18.4 | 2951 | 10 | US-10-450-763-21970 Sequence 35, Appl |
| 35 | 689 | 18.3 | 1705 | 8 | US-10-479-334-35 Sequence 21, Appl |
| 36 | 645.5 | 17.2 | 1670 | 8 | US-10-479-334-21 Sequence 7, Appl1 |
| 37 | 640 | 17.0 | 1521 | 8 | US-10-479-334-7 Sequence 61, Appl |
| 38 | 636 | 16.9 | 522 | 5 | US-09-720-086-61 Sequence 61, Appl |
| 39 | 636 | 16.9 | 522 | 9 | US-10-623-813-61 Sequence 199022, Sequence 199023, |
| 40 | 629 | 16.7 | 588 | 6 | US-10-027-632-199022 Sequence 199023, |
| 41 | 629 | 16.7 | 588 | 6 | US-10-027-632-199023 Sequence 199022, |
| 42 | 629 | 16.7 | 588 | 7 | US-10-027-632-199022 Sequence 199023, |
| 43 | 629 | 16.7 | 588 | 7 | US-10-027-632-199023 Sequence 18, Appl |
| 44 | 609 | 16.2 | 435 | 5 | US-09-720-086-18 Sequence 18, Appl |
| 45 | 609 | 16.2 | 435 | 9 | US-10-623-813-18 Sequence 18, Appl |

ALIGNMENTS

RESULT 1
US-10-623-813-84
Sequence 84, Application US/10623813
Publication No. US20040234997A1
GENERAL INFORMATION:
APPLICANT: Li, En
APPLICANT: Okano, Masaki
APPLICANT: Xie, Shaping
TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
FILE REFERENCE: 0609.456003
CURRENT APPLICATION NUMBER: US/10/623, 813
CURRENT FILING DATE: 2003-07-22
PRIOR APPLICATION NUMBER: US 09/720, 086
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: PCT/US99/14373
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/090, 906
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: US 60/093, 993
PRIOR FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 119
SOFTWARE: PatentIn version 3.2
SEQ ID NO 84

LENGTH: 2371
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-623-813-84

Alignment Scores:

| Pred. No.: | 0 | Length: | 2371 |
|------------------------|---------|---------------|------|
| Score: | 3763.00 | Matches: | 689 |
| Percent Similarity: | 100.0% | Conservative: | 0 |
| Best Local Similarity: | 100.0% | Mismatches: | 0 |
| Query Match: | 100.0% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-10-623-813-86 (1-689) x US-10-623-813-84 (1-2371)

```

Oy      1 Metasna1aValGluGluAsnGlnGlyProGlyGluSerGlnuValGluGluAsnSer 20
Db      217 ATGAATGCTGTGGAGAAAGAAACCAAGGGGCCCGGGAGTCTCAGAAAGTGGAGAGCCAGC 276
Oy      21 ProProa1aValGlnGlnProThrAspProAlaSerProThrValAlaThrThProGlu 40
Db      277 CTTCTGCTGTGACAGCAAGCCCACTGACCCGCACTCCCACTGTGCTACCAAGCCTAG 336
Oy      41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
Db      337 CCGGTGGGGTCCGATGCTGGGGACAAAGATGCAACCAAGCAGGGGATGACGAGCCAGAG 396
Oy      61 TyriGluAspGlyAArgGlyPheGlyY1IleGlyLuleuValITTPGlyLysLeuArgGlyPhe 80
Db      397 TACAGAGCAGCGCCGGGCTTTGGCATTTGGGAGCTGTGTGGGGAAACTCGGGGGCTTC 456
Oy      81 SerTrpTrpProGlyAArgGlyLeuValSerTrpTrpMetThrGlyAArgSerArgAlaGlu 100
Db      457 TCCGTGTGGCCAGGGCCGCAATGTGTCTTGATGATGAGAGGGCCGAGCCAGCACTGAA 516
Oy      101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValAlaCysValGluLys 120
Db      517 GGCAACCGCGTGGTCAATGTGGTGGAGACGCAAAATCTCGATGTGTGTGATGAGAAG 576
Oy      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyraAsnLysGlnPro 140
Db      577 CTGATGCCGCTGAGCTCTTTTTCAGTGCCTGCCACAGGCCCACTAACCAAGCAGGCC 636
Oy      141 MetTyraArgLysAlaAlaIleTyriGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      637 ATGTAACCAAGCAAGCATTAAGAGGTCTTGACAGTGGCCAGCAGCCGCGGGAGAGCTG 696
Oy      161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValAsn 180
Db      697 TTCCCGGTGTCCACGACAGCATGAGTACACTGCAAGGCCCTGTGAGGTGCAGAAC 756
Oy      181 LysProMetI1IleGlyTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      757 AAGCCCAAGATTGATGAGGCTCTGGGGGCTTCACGCTTCTGCGCCCTTAAGGGCCTGAG 816
Oy      201 ProProGluGluGluLysAsnProTyriLysGluValITyThrAspMetTrpValGluPro 220
Db      817 CCACCAAGAAAGAAAGAAAGATCTCTTACAAAGAGTGTACAGGACATGTGGGTGAACCT 876
Oy      221 GluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 240
Db      877 GAGGAGCTGCTACGACCACTCCACACGCAAAAGCCCCGGAAGACGACAGCGAG 936
Oy      241 LysProLysValLysGluL1IleL1eAspGluArgThrArgGluArgLeuValITyriGluVal 260
Db      937 AAGCCCAAGGCTCAAGGAGATTATGATAGCGCACAAAGACCGGCTGTGTACAGAGTG 996
Oy      261 ArgGlnLysCysArgAsnI1IleGluAspI1IleCysI1IleSerCysGlySerLeuAsnValThr 280
Db      997 CGGCAAGAGTCCGGAACATTGAGACATCTGCACTCTCTGTGGAGGCTCAATTTACC 1056
Oy      281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
  
```

```

Db      1057 CTGGAACACCCCTCTGTGTGAGAGATGTGCCAAACTGCAGAACTGCTTCTGAG 1116
Oy      301 CysAlaTyriGlnITyraAspAspGlyTyriGlnSerTyriCysThriLysCysGlyGly 320
Db      1117 TGTGCGTACCAAGTACAGACAGAGAGCGGTACCGATCTTCACTGACACATCTGCTGTGGGGCC 1176
Oy      321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGlyCysVal 340
Db      1177 CGTAGAGTCTCACTGTGCGGAACAACAACAACACTGCTGAGGTCTTTGTGGAGGTGTG 1236
Oy      341 AspLeuLeuValGlyProGlyValAlaGlnAlaAlaI1IleLysGluAspProITyPAsnCys 360
Db      1237 GACCTTGTGTGGGGCCGGGGGCTCCAGAGCACCATTAAAGAAAGCCCTGGAACCTGC 1296
Oy      361 TyriMetCysGlyHisLysGlyTyThrTyriGlyLeuLeuArgArgGluAspTrpSer 380
Db      1297 TACATGTGGGACAAAGGTACTTACGGGTCTGTGGCGGCGAGAGACTGGCCCTCC 1356
Oy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValITyr 400
Db      1357 CGGCTCAAGATGTTCTTCGTAATTAACAGACAGGAATTGACCTCCAAAGTTTAC 1416
Oy      401 ProProValProAlaGluLysArgLysProI1IleArgValLeuSerLeuPheAspGlyIle 420
Db      1417 CCACCTGTCCAGCTGAAGAGAGAGAGCCCATCGGGTGTGTCTCTTTGATGAGAAATC 1476
Oy      421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyI1IleGlnValAlaAspArgTyriAlaAsn 440
Db      1477 GCTTACAGGGCTCTGTGTGTGTAAGACTTGGGCACTTCAAGTGGACCGGTACATTCCTCG 1536
Oy      441 GluValCysGluAspSerI1IleThValGlyMetValArgHisGlnGlyLysI1IleMetTyr 460
Db      1537 GAGGTGTGTAGGACTCATACAGGTGGCATGTGTGGCCAGCAGGAGAAATCATGTAC 1596
Oy      461 ValGlyAspValAArgSerValThriGlnLysHisI1IleGlnGluITTPGlyProPheAspLeu 480
Db      1597 GTCGGGAGCGTCCCAAGCTCACAGAGAGATATCCAGAGGTGGGGCCATTGCATCTG 1656
Oy      481 ValI1IleGlyLysProCysAsnAspLeuSerI1IleValAsnProAlaArgLysGlyLeu 500
Db      1657 GTGATTTGGGGCAATCCCTGCAATGACTCTTCCATCTCTCAACCTGCTCGAAGGGCTTC 1716
Oy      501 TyriGlnGlyThriArgLeuPhePheGluPheTyriArgLeuLeuHisAspAlaArgPro 520
Db      1717 TACAGGGCACTGGCCGAGCTCTTCTTGAGTTTACCGGCTCTGATGATGAGCCGGGCC 1776
Oy      521 LysGluGluAspAspAArgProPhePheTrpLeuPheGluAsnValAlaAlaMetGlyVal 540
Db      1777 AAGAGGAGATGATCGCCCTTCTTGTGCTTTTGAAGATGTGTGGCCATGGGCGCTT 1836
Oy      541 SerAspLysArgAspI1IleSerArgPheLeuGluLeuSerAsnProValMetI1IleAspAlaLys 560
Db      1837 AGTACAAAGAGGACATCTGCGATTTCTGAGTCCAAACCTGTGATGATGATGCCAA 1896
Oy      561 GluValSerAlaAlaHisArgAlaArgTyriPheITPGLysLeuLeuProGlyMetAsnArg 580
Db      1897 GAATGTCAAGTGCACACAGAGGCCGCTACTTGGGGTAACTTCCCGGTATGAACAGG 1956
Oy      581 ProLeuAlaSerThriValAsnAspLysLeuGluLeuGlnGluCysLeuGluHisGlyArg 600
Db      1957 CCGTTGGCATTCACCTGAATGATGATGAGTGAAGCTGACAGAGGTCTGAGACATGGGCGG 2016
Oy      601 I1IleAlaLysPheSerLysValArgThriI1IleThThriArgSerAsnSerI1IleLysGlnGly 620
Db      2017 ATAACCAAGTTCCAGCAAGTACAGACCATTAATCTGAGAGTCAAACTCATTAAGCAGGGC 2076
Oy      621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspI1IleLeuITyPTrpCysThriGlu 640
Db      2077 AAAGACAGCAATTTCTCTGCTTCATGATGAGAAAGAGGACATCTTATGTGTGACCTGAA 2136
Oy      641 MetGluArgValPheGlyPheProValHisTyriThraAspValSerAsnMetSerArgLeu 660
Db      2137 ATGAAAGGGATTTGTGTTCCAGTCACTATCTAGTACGTTCCAACTCATAGAGCGGCTTG 2196
  
```

QY 661 AlArGlnAArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2197 GCGAGGCGAGAGACTGCTGGGCGGCTCAAGGCGGTGCGACATCCGCCACTTCTGGCT 2256
QY 681 ProLeuYsgLurYrPheAlaCySVal 689
Db 2257 CCGCTGAAGAGATATTTTGGCTGTGTG 2283

RESULT 2
US-10-144-577-2
; Sequence 2, Application US/10144577
; Publication No. US20030083292A1
GENERAL INFORMATION:
; APPLICANT: MacLeod, Alan Robert
; TITLE OF INVENTION: Inhibitors of DNA Methyltransferase Isoforms
; FILE REFERENCE: MET-005
; CURRENT APPLICATION NUMBER: US/10/144,577
; PRIORITY FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,202
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,212
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3005
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-577-2

Alignment Scores:
Pred. No.: 0 Length: 3005
Score: 3763.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0

US-10-623-813-86 (1-689) x US-10-144-577-2 (1-3005)

QY 1 MetAsnAlaValGluGluAsnGlnGlyProGlyGluSerGlnValGluGluAlaSer 20
Db 899 ATAAATGCTGTGAGAGAAACCAAGGCGCGGAGCTCTCAAGAGTGGAGAGGCCAC 958
QY 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db 959 CCTCTGCTGTGAGAGCCCACTGACCCCGCATCCCTGCTGCTACACGCTGAG 1018
QY 41 ProValGlySerAspAlaGlyAspValAspAlaThrValAlaGlyAspAspGluProGlu 60
Db 1019 CCGGTGGGCTCCATGCTGGGGACAAATGCCCAAGCGAGCGATGACAGCCAGAG 1078
QY 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
Db 1079 TACGAGGACGGCGGGGCTTTGGCATTTGGGAGCTGTGTGGGGAACCTGGCGCTTC 1138
QY 81 SerTrpTrpProGlyArgGlyLeuValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db 1139 TCTGTGGGCGCAGGCGCATTTGTCTTGGTGTGATGACGGGCGCGAGACCGAGCTGAA 1198
QY 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
Db 1199 GGCACCGGCTGGGTCAATGTGAGTTCGAGAGCGCAAAATTCACATGTGTGTGAGAG 1258
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db 1259 CTATATGCGCTGAGCTGTCTTTGCAAGTCTTCCACAGGCGCATGTCACAAGCACCCC 1318
QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db 1319 ATGTACCGCAAAACCATCTACAGAGTCTCTGACAGGTGACAGACCGCGGGGAACCTG 1378

QY 161 PheProValCysHisAspSerAspGluSerAspThrAlaValAlaGluValGlnAsn 180
Db 1379 TTCCCGGTGTGCCACGACGACGATGAGAGTGAACACTGCCAAGCCCGTGGAGGTGCACAC 1438
QY 181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db 1439 AACCCCATGATTAATGGGCGCTGGGGGGCTTCAGCTTCTGCGCTTCAAGGGGCTGGAG 1498
QY 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db 1499 CCACCAAGAAAGAAAGAAATCCCTCAAGAAAGTGAACAGGACATGTGGTGGAACT 1558
QY 221 GluAlaAlaAlaTyrAlaProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db 1559 GAGGAGCTGCTTACCGACCACTTCCACAGCCCAAAAGCCCGGGAAGACACAGCGGAG 1618
QY 241 LysProLysValLysGluIleLeuAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db 1619 AACCCCAAGGTCAAGGAGATTATGTGAGCGCACAAAGAGCGGCTGTGTACGAGGTG 1678
QY 261 ArgGluLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 1679 GCGCAGAAAGTCCGGAACATTGAGACATCTGCATCTCTGTGGAGCTCAATGTATACC 1738
QY 281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCySphLeuGlu 300
Db 1739 CTGGAACACCCCTCTTCTGTGAGAGAAATGTGCCAAACATCGCAAGAACTGCTTCTGAG 1798
QY 301 CysAlaTyrGlnTyrAspAspArgGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1799 TGTGCTTACCAATGACGACGACGAGCTTACGCTTACGACATCTCTGCTGGGGG 1858
QY 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1859 GGTGAGGTGCTCATGTGGGAAACAACATGCTGAGGCTTTTGGAGTGTGTG 1918
QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTyrAsnCyS 360
Db 1919 GACCTCTGGTGGGCGGGGGCTGCGCAGGCGCATTTAAGAAAGCCCTGGAACCTGC 1978
QY 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgAlaGluAspTrpProSer 380
Db 1979 TACATGTGCGGACAAAGGTACTTACGGGCTGCTCGGGGAGAGGACTGGCCCTCC 2038
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
Db 2039 CGGCTCCAGATGTTCTTCCCTAATAACACAGCACAGAAATTTCACCTCCAAAGTTTAC 2098
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 2099 CACCTGTCCAGCTGAGAAAGAAAGCCCATCGGGTGTCTCTTGTGATGGAATC 2158
QY 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 2159 GCTACAGGGCTCTGTGTGTAAGGACTTGGGCATTCAGGTGACCGCTCATATGGACTCG 2218
QY 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 2219 GAGGTGTGAGGACCTCATCACGTTGGGATGTGGGACCAAGGGAAGATCATATAC 2278
QY 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db 2279 GTGCGGAGCTCCGACGCTGACACAGAAAGCATATCAGAGAGTGGGGCCATTGCACTG 2338
QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db 2339 GTGATTTGGGGGAGTCCCTGCAATGACCTTCATGTCAACCTGCTCCCAAGGGCTC 2398
QY 501 TyrGluGluTyrThrLysArgLeuPhePheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 2399 TACGAGGACCTGGCGGCTCTTTGAGTTTACCGGCTCTGCAATGATGGCGGCCCC 2458
QY 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540

Db 2459 AAGGAGGAGATGATCCCTTTCTTCTGCTCTTTAGATGATGCGCATGGCGCTT 2518
Qy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 2519 AGTGACAAAGGAGCATCTGGCATTTCTCGAGTCCAAACCTGTGATGATGATGCCAA 2578
Qy 561 GluValSerAlaAlaIleArgAlaArgTyrPheTrpGluYasnLeuProGluMetAsnArg 580
Db 2579 GAAGTGTCACTGCACACAGGCGCGCTACTTGGGGTAACTTCCCGGTATGAACAGG 2638
Qy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGluCysLeuGluHisGlyArg 600
Db 2639 CCGTTGGCATCTCACTGTGAATGATAAAGCTGAGCTGCAGAGGTCTTCGAGCATGGCAGG 2698
Qy 601 IleAlaLysPheSerLysValArgThrIlePheThrArgSerAsnSerIleLysGlnGly 620
Db 2699 ATAGCCAAAGTTCAACAAAGTAGACCATTTACTACGAGGTCAAACTCCATTAACAGGGC 2758
Qy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db 2759 AAGAGCCAGCATTTCTGTCTTATGATAGAAAGAGGACATCTATGGTGCACGTAA 2818
Qy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2819 ATGGAAAGGATATTGGTTTCCAGTCCACTATTAAGCTGCTCCAAACATGAGCCGCTTG 2878
Qy 661 AlaArgGlnArgLeuLeuGlyArgSerTyrSerValProValIleArgHisLeuPheAla 680
Db 2879 GCGAAGGAGAGACGTCTGGCGGCTCATGGAGCGTGCAGCATCCGCACTCTTCGCT 2938
Qy 681 ProLeuLysGluTyrPheAlaCysVal 689
Db 2939 CCGCTGAAGAGTATTGTGCGTGTGTG 2965

RESULT 3

US-10-172-118-132
/ Sequence 132, Application US/10172118
/ Publication No. US20030224374A1
/ GENERAL INFORMATION:
/ APPLICANT: Dai, Hongyue
/ APPLICANT: He, Yundong
/ APPLICANT: Linsley, Peter
/ APPLICANT: Mao, Mao
/ APPLICANT: Roberts, Chris
/ APPLICANT: Van 't Veer, Laura
/ APPLICANT: Van de Vijver, Marc
/ APPLICANT: Bernards, Rene
/ TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
/ FILE REFERENCE: 9301-175-999
/ CURRENT APPLICATION NUMBER: US/10/172,118
/ CURRENT FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 60/380,770
/ PRIOR FILING DATE: 2002-05-14
/ NUMBER OF SEQ ID NOS: 2699
/ SEQ ID NO 132
/ LENGTH: 3005
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: AF067972
/ DATABASE ENTRY DATE: 2001-06-18
/ US-10-172-118-132

Alignment Scores:

Pred. No.: 0 Length: 3005
Score: 3763.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-623-813-86 (1-689) X US-10-172-118-132 (1-3005)

Qy 1 MetAsnAlaValGluGluAsnGlnGlyProGluGlySerGlnLysValGluGluAlaSer 20
Db 899 ATAAATGCTGTGAGAAAGAAACAGAGGCGCCGGGAGTCTCAAGAGGTGAGAGGCCAGC 958
Qy 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db 959 CTTCTGCTGTGACAGAGCCCACTGAACCCCGCATCCCGCACTGTGGCTACAGAGCCCTGAG 1018
Qy 41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaLysAspAspGluProGlu 60
Db 1019 CCGGTGGGCTCCATCTGTGGGAGCAAGAAAGCCAAACAGAGCCCATGACAGCCAGAG 1078
Qy 61 TyrGluAspGlyValArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
Db 1079 TAGAGAGCCGCGCGGCGCTTTGGCATTTGGGAGCTGTGTGGGGAAATCCGGGCGCTTC 1138
Qy 81 SerTrpTrpProGlyValArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db 1139 TCTGGTGGCCAGGCGGCATTTGTCTTGGTGATGACGGGCGGAGCCGAGAGCTGAA 1198
Qy 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
Db 1199 GGCACCCGCTGGGTCAATGTGGTTCGAGAGAGCGCAAAATCTCAAGTGTGTGTGAGAG 1258
Qy 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db 1259 CTAGTCCGCTGAGCTGCTTTTTCGATGCTTCCACAGGCCACGTAACAAAGCAACCCC 1318
Qy 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db 1319 ATGACCGGCAAGACCAATCTACAGAGTCTTCAGAGTGGCCAGAGCCGCGGGGAACTG 1378
Qy 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysValAlaGluValGluAsn 180
Db 1379 TTCCCGGTGTGCACACAGCAGATGAGAGTGCATCTCCAAAGCCGAGGAGTGCAGAAC 1438
Qy 181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db 1439 AAGCCCATGATTAAATGGGCTTGGGGGCTTTCAGGCTTCTGGCCTTAAGGGCTTGAG 1498
Qy 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db 1499 CCAACGAAAGAGAAAGAAATCCCTCAAGAAAGTGAACAGACATGTGGTGAACCT 1558
Qy 221 GluAlaAlaAlaTyrAlaProProProProAlaLysSerProArgLysSerThrAlaGly 240
Db 1559 GAGGAGCTCTTACGACCACTCCACCAAGCCCAAAAGCCCGGAAGACACAGCGAG 1618
Qy 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db 1619 AAGCCCAAGGTCAAGGATTAATTGATGACGCAAGAGAGCGGTGCTGACAGAGTG 1678
Qy 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 1679 CGCGAGAGTGGCGGAACATTGAGGACATCTGCATCTCTGTGGAGCTCAATGTTACC 1728
Qy 281 LeuGlnHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db 1739 CTGAAACACCCCTCTTCTGTGGAGGATGTGCCAAATCTGCAAGAACTCTTCTTGAG 1798
Qy 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db 1799 TGTGCGTACAGTACAGACAGACAGCGGTACAGTCTCACTGACCACTCTGTGGGGGC 1858
Qy 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1859 CGTAGAGTGTCTCATGTGCGGAAACAAACACTGCTGAGGTGCTTTTCCGTGGAGTGTG 1918
Qy 341 AspLeuLeuValGlyProGlyValaGlnAlaAlaIleLysGluAspProTrpAsnCys 360
Db 1919 GACCTTGTGTGGGCGCGGGGCTGCCAGGACCATTAAGAGAACCCCTGGAACCTGC 1978

```

Oy 361 TyMeCvSeGlyHslySGlyThrTyrGlyLeuLeuAraGArGluAspTrpProSer 380
| | | | |
Db 1979 TACATGTCGGGGCCAAAGGGTCACTACGGCTGCTCGGGCGGCGAGAGACTGGCCCTCC 2038
Oy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
| | | | |
Db 2039 CGGCTCCAGATGTTCTTCGCTAATTAACACAGACACAGAAATTGACCTCCAAAGATTAC 2098
Oy 401 ProProValProAlaGluValArgLysProLysLeuArgValLeuSerLeuPheAspGlyTle 420
| | | | |
Db 2099 CCACCTGTCCACGCTGAGAGAGAGAACCCCATCCGGGGCTGTCTCTCTTGAATGGAATC 2158
Oy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyTleGlnValAspArgTyrTleAlaSer 440
| | | | |
Db 2159 GCTACAGAGGCTCTGTGCTGAGAGACTTGGGCAATTCAGGTGACCCGCTCATTTGCTCG 2218
Oy 441 GluValCysGluAspSerLethrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
| | | | |
Db 2219 GAGGTGTGAGAGCATCATCAAGCTGGGCAATGGTGGGCGCACGGGGAAGATCATGTAC 2278
Oy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
| | | | |
Db 2279 GTCGGGAGCCTCCGACAGCTCACAGAGCATATCCAGAGTGGGGCCCATTCGATCTG 2338
Oy 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
| | | | |
Db 2339 GTATATGGGGGCGAGTCCCTCGAATGACTCTTCACATCCGCAACCTTGCTCGAAGGGCTC 2398
Oy 501 TyrGluGlyThrGlyArgLeuPheGlnPheTyrArgLeuLeuHisAspAlaArgPro 520
| | | | |
Db 2399 TAGAGGGGCACTGGCCGGCTCTTCTTGAGTTCAACGGCTCCCTCGCATGATGGCGGCC 2458
Oy 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValValAlaMetGlyVal 540
| | | | |
Db 2459 AAGGAGGGAATGATCCGCCCTTCTTGCTGCTTGAATGATGGTGGCCATGGGGCTT 2518
Oy 541 SerAspLysArgAspLysSerArgPheLeuGluSerAsnProValMetCileAspAlaLys 560
| | | | |
Db 2519 AGTGACAGAGGGGCACTCGCGCATTTCTCGAGTCAACCCCTGATGATTTGATGCCAA 2578
Oy 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
| | | | |
Db 2579 GAATGTCACCTGCACACAGGGCCGCTACTTCTGGGGTAACTTCCCGGATGAACAGG 2638
Oy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGlyCysLeuGlnHisGlyArg 600
| | | | |
Db 2639 CCGTTGGCATCCACTGTGAATGATTAAGCTGAGAGTGTCTGAGAGCAAGGACAGG 2698
Oy 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
| | | | |
Db 2699 ATAGCCCAAGTTCAGCAAGTAGGAGCACTTAAGAGAGTCAAGCTCCATTAAGCAGGGC 2758
Oy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspLysLeuTyrPyrGlu 640
| | | | |
Db 2759 AAAGACACAGACTTTCTGCTTTCATGATGAGAAAGAGAGCATCTTATGTTGACAGAA 2818
Oy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
| | | | |
Db 2819 ATGGAAGGGGTATTGGTTTCCAGTCCACTATATCTGACGCTCCCAACATGAGCGGCTTG 2878
Oy 661 AlaArgGlnArgLeuGluArgSerTrpSerValProValIleArgHisLeuPheAla 680
| | | | |
Db 2879 GCGAGGAGAGACTGCTGGGCGGCTCATGAGCGGTGACATCCGACCTCTTGCT 2938
Oy 681 ProLeuLysGluTyrPheAlaCysVal 689
| | | | |
Db 2939 CCGCTGAAGAGTATTTTGCCTGTGTG 2965

```

RESULT 4
 US-10-342-887-132
 ; Sequence 132, Application US/10342887
 ; Publication No. US20040058340A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Robertes, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 132
; LENGTH: 3005
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-132

Alignment Scores:
Pred. No.: 0 Length: 3005
Score: 3763.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0

US-10-623-813-86 (1-689) x US-10-342-887-132 (1-3005)
Oy 1 MetAsnAlaValGluGluAsnGlnGlyProGlyLysSerGlnLysValGluGluAlaSer 20
| | | | |
Db 899 ATGAATGCTGTGAAGAAACCAAGGGGCGGGGAGTCTCAAGAGGTGAGAGGAGGACGC 958
Oy 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrTrpProGlu 40
| | | | |
Db 959 CCTCTGCTGTGACAGAGCCCACTGACCCCGCATCCCTGCTGCTACCAAGCCCTGAG 1018
Oy 41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
| | | | |
Db 1019 CCGTGGGCTCCAGTCTGGGGGCAAGATGCCAACAGGCGGATGACAGCCAGAG 1078
Oy 61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
| | | | |
Db 1079 TAGAGAGAGCGGGGGCTTTGGCATTTGGGAGCTGGTGTGGGGAAACTGGGGGCTTC 1138
Oy 81 SerTrpTrpProGlyArgIleValSerTrpTrpMetThrGlyArgSerArgAlaAlaGlu 100
| | | | |
Db 1139 TCTGTGTGGCCAGAGCCGCAATTTGTGTCTGTGTGATGACGGGCGGAGCCGAGCAGCTGAA 1198
Oy 101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
| | | | |
Db 1199 GGCACCGGCTGGGTCATGTGTGTGGAGAGCGCAATTTCAAGTGTGTGTGTAAG 1258
Oy 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
| | | | |
Db 1259 CTGATGCGCTGAGCTCGTTTTCAGTGGCTTCACAGGCGCAGAGCAAGCAAGCAGCC 1318
Oy 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
| | | | |
Db 1319 ATGTACCGCAAAAGCATCTACGAGTCTTCAGAGTGGCCAGAGCCGCGGGGAAACTG 1378
Oy 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValAsn 180
| | | | |
Db 1379 TTCCCGGTGTGCCACAGACGATGAGAGTGAACACTGCCAAGCCGCTGAGGTGCABAAC 1438
Oy 181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
| | | | |

```



```

Db      1439 AAGCCATGATTGATGGGCCCTTGGGGGCTTCCAGCCTTCTGCGCTTAAGGCGCTGGAG 1498
Qy      201 ProProGluGluGluLeuAsnProTyrIleGluValIleThrAspMetTyrValGluPro 220
Db      1499 CCACGAGAGAGAGAGAGATCTCTCAAGAGAGTACACGAGCATGTGGTGGAACT 1558
Qy      221 GluAlaAlaAlaTyrAlaProProProAlaValLeuValProAlaGlySerThrIleGlu 240
Db      1559 GAGGAGCTGCTTACGACACCTCCAGCCAAAAGCCCCGGAGAGGACACAGCGAG 1618
Qy      241 LysProLysValIleGluIleIleAspGluArgThrArgGluArgLeuValIleGluVal 260
Db      1619 AAGCCAAAGTCAAGAGATTATGATGAGCCACAGAGAGCGCTGTGTAGCAGGTG 1678
Qy      261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1679 CGGAGAAAGTCCGGAAATTGAGGACATCTGCATCTCTGTGGAGGCTCATTTTACC 1738
Qy      281 LeuGlnHisProLeuPheValIleGlyMetCysGlnAsnCysValAsnCysPheLeuGlu 300
Db      1739 CTGGAAACCCCCCTTCTGTGGAGAAAGTGCCAAAAGTGCAGAACCTGCTTTGGAG 1798
Qy      301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1799 TGTGCGTACCAAGTACGACGAGCGGCTACCAAGTCTTACCATCATCTGTGTGGGGC 1858
Qy      321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValIleCysVal 340
Db      1859 CGTAGAGTCTCATGTGCGGAAACAACTGCTCAGAGTGTGTTGGCTGAGAGTGTG 1918
Qy      341 AspLeuLeuValIleGlyProGlyValAlaAlaAlaIleValGluAspProTyrAsnCys 360
Db      1919 GACCTCTTGAGGGGCGGGGGCTCCAGCGAGCCATTAAAGAAAGCCCTTGAACTGC 1978
Qy      361 TyrMetCysGlyHisIleGlyThrTyrGlyLeuLeuArgArgGluAspTyrProSer 380
Db      1979 TACATGTGCGGGCACAAGGATCTTACGGGCTGCGGGCGGCGAGAGACTGGCGCTCC 2038
Qy      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysAlaTyr 400
Db      2039 CGGCTCCAGATGTTCTTGGCTAAATCAACGACGCAATTTGACCTCCAAAGGTTTAC 2098
Qy      401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2099 CCACCTGTCCAGCTGAGAAAGAGAGAGGCCATCCGGGTGCTGTCTCTTTGATGATC 2158
Qy      421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2159 GCTACAGGGGCTCTGTGCTCTGAAGACTTGGGCATTCAAGTGGACCGCTACATTGCTCG 2218
Qy      441 GluValCysGluAspSerIleThrValIleValIleValIleGlnGlyIleMetTyr 460
Db      2219 GAGGTGTGTGAGATCTCATACGGTGGCATGTGGCGGACCAAGGGAATCATCTGTAC 2278
Qy      461 ValGlyAspValIleGlySerValIleThrGlnLysHisIleGlnGluTyrGlyProPheAspLeu 480
Db      2279 GTCCGGAGACCTCCGAGCGCTACACAGAAACATATCCAGAGTGGGCCCATTCATCTG 2338
Qy      481 ValIleGlyIleGlySerProCysAsnAspLeuSerIleValIleAspProAlaArgLysGlyLeu 500
Db      2339 GTGATTTGGGGGAGTCCCTGTCAATGACTTTCATCGTCAACCTGCTCGCAAGGGCTCC 2398
Qy      501 TyrGlnGlyThrArgLysArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db      2399 TACAGGGGCACTGGCGGCTCTTCTTGAAGTTTCAACCCCTCTCGCAAGAGTGGCGGCC 2458
Qy      521 LysGlnGlyAspAspArgProPhePheThrLeuPheGlnAsnValIleAlaMetGlyVal 540
Db      2459 AAGAGGAGAGATGATCGCCCTTCTTCTGCGCTCTTGAAGATGTGTGGCCATGGCGCTT 2518
Qy      541 SerAspLysArgAspIleSerArgPheLeuGlnSerAspProValMetIleAspAlaLys 560
Db      2519 AGTGACAAAGAGGACATCTCGGATTTCTCGAGTCCAACCTGTGATGATTTGATGCCAAA 2578

```

```

Qy      561 GluValSerAlaAlaHisArgAlaArgTyrPheThrGlyValLeuProGlyMetAsnArg 580
Db      2579 GAATGTCTACCTGACACAGAGCCCGGCTACTTCTGGGGTAACTTCCGGATATGAAACAG 2638
Qy      581 ProLeuAlaSerThrValAsnAspLysLeuGlnIleGlnGlyCysLeuGlnHisGlyArg 600
Db      2639 CCGTTGGCATCTGTGAATGATTAACCTGGAGCTGCAGAGAGTCTTGGAGCATGGCAGG 2698
Qy      601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleGlyGlnGly 620
Db      2699 ATACCCAAAGTTCAACAAAGTACGACCATTAATACGAGGTCAAACTCATTAACACAGGC 2758
Qy      621 LysAspGlnHisPheProValPheMetAsnGlnLysGlnAspIleLeuTyrCysThrGlu 640
Db      2759 AAAGACCAAGATTTTCTGCTTCTATGAAGAAGAGACATCTTAAGTGCATGAA 2818
Qy      641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2819 ATGGAAGGCGTATTTGGTTTCCAGTCCAGTCACTGACGTCTCCAAACATAGCGCGCTTG 2878
Qy      661 AlaArgGlnArgLeuLeuGlnValArgSerTyrPheSerValProValIleAlaGlnLeuPheAla 680
Db      2879 GCGAGGAGAGACTGTGCGCGGCTCATGAGCGTGCACATCATCCGACCTTCTGCT 2938
Qy      681 ProLeuLysGluTyrPheAlaCysVal 689
Db      2939 CCGCTGAAGAGATTTTGGCTGTGTG 2965

RESULT 5
US-10-330-773-279
; Sequence 279, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 279
; LENGTH: 3009
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-279

Alignment Scores:
Pred. No.: 0 Length: 3009
Score: 3763.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-623-813-86 (1-689) x US-10-330-773-279 (1-3009)
Qy      1 MetAsnAlaValGluGlnGlnGlyProGlyGlnSerGlnLysValIleGluAlaSer 20
Db      903 ATGAATCTGTGGAAAGAAACAGAGGGGCCGGGGAGTCTCGAAGGTGGAGAGCGCCAGC 962
Qy      21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      963 CCTCTCTGTGGAGAGCCACTGACCCGACATCCCACTGTGGCTTACCAAGCTGTAG 1022
Qy      41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
Db      1023 CCGTGGGGTCCGATGTGGGGACAAAGATGCCCAAGAGCGATGACGAGCCAGAG 1082
Qy      61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
Db      1083 TACAGAGACCGCCGGGCTTTGGCATTTGGGGAAGCTGTGTGTGGGGAACCTGCGGGGCTTC 1142

```

QY 81 SerTrpTrpProGluValArgIleValSerTrpTrpMetThrGluArgSerArgAlaIleGlu 100
DB 1143 TCCTGGTGGCCAGGCGCATGTGTCTTGTTGGATGAGCGGCGGAGCGGACACTGAA 1202
QY 101 GluThrArgTrpValMetTrpPheGluValysPheSerValValCysValGluLys 120
DB 1203 GGCAACCGCTGGGTCTATGTGGTTCGGAGACGGCAAAATCTCAGTGTGTGTCTTGAGAA 1262
QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTrpAsnLysGlnPro 140
DB 1263 CTGATGCGCGCTGAGCTCTTTTGTGACGTCCGTCACAGGCCACGTACACAGCAGGCC 1322
QY 141 MetTrpArgLysAlaIleTrpGluValLeuGlnValAlaSerSerArgAlaGluLysLeu 160
DB 1323 ATGTACCGCAAGACCATCTACAGGCTCTGACAGTGGCGACAGCGCGCGGAGAGCTG 1382
QY 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
DB 1383 TTCCCGGTGTCCACGACAGCGATGAGTGCACCTGCCAAGCCGTGGAGGTGCAGAAC 1442
QY 181 LysProMetIleGluTrpAlaLeuGluGluPheGlnProSerGluProLysGluLeu 200
DB 1443 AAGCCCATGATTGAATGGGCGCTGGGGGCTTCCAGCCTTCTGGCCCTAAGGGCGTAG 1502
QY 201 ProProGluGluGluLysAsnProTrpTrpLysGluValTrpThrAspMetTrpValGluPro 220
DB 1503 CCACCAAGAAAGAAAGAAATCCCTCAAAAGAAAGTACAGCGACATGTGGTGGAACT 1562
QY 221 GluAlaAlaAlaTrpAlaProProProAlaLysLysProAlaGluSerThrAlaGlu 240
DB 1563 GAGGCACTGCTGACGACCACTCCACCAAGCAAAAGCCCCGGAAGAGCAGCGGAG 1622
QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTrpGluVal 260
DB 1623 AAGCCCAAGGTCAAGGATTAATGATGACGCCACAGAGAGCGGCTGTGTACGAGGTG 1682
QY 261 ArgGlnLysCysAspAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1683 CGGAGAAAGTCCGGAACATTTGAGAGACATCTGCACTCTGTGGAGGCTTCAAGTTTACC 1742
QY 281 LeuGlnHisProLeuPheValGluLysCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 1743 CTGGAACACCCCTCTTCTGTGGAGGAATGTGCCAAACTCCAGAACTGCTTTCTGGAG 1802
QY 301 CysAlaTrpGlnTrpAspAspAspGluTrpGlnSerTrpCysThrIleCysCysGluGly 320
DB 1803 TGTGCGTACAGTACGACGACGAGCGGCTACCAAGTCTACTGCACATCTGCTGTGGGGC 1862
QY 321 ArgGluValLeuMetCysGluValAsnAsnAsnCysCysAspArgCysPheCysValGluCysVal 340
DB 1863 CGTGAGGTGCTCATGTGCGGAAACAAACAATGCTGCAGGTCTTTTGTGGAGTGTGTG 1922
QY 341 AspLeuLeuValGluProGluValAlaGlnAlaIleLysGluAspProTrpAsnCys 360
DB 1923 GACCTCTTGGTGGGGCGGGGGCTGCCAGGCACTTAAGAAAGACCTTGGAACTGC 1982
QY 361 TyrMetCysGluHisLysGluTrpTrpGluLeuValArgArgGluAspTrpProSer 380
DB 1983 TACATGTGGCGGCAACAAGGTACTTACGGGCTGCTGGCGCGGCGAGAGGACTGGCCCTCC 2042
QY 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
DB 2043 CGGCTCCAGAGTCTTCTGCTTAATTAACAACAAGAAATTTGACCTTCCAAAGGTTTAC 2102
QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGluLe 420
DB 2103 CCACCTGTCCAGCTGAGAGAGAGAACCCATCCGGGTGCTGTCTTCTTTGATGAGATC 2162
QY 421 AlaThrGluLeuLeuValLeuLysAspLeuGluIleGlnValAspArgTrpIleAlaSer 440
DB 2163 GCTACAGAGGCTCCTGTGTGTAAGAGACTTGGGCACTTCCAGGTGAGACCGCTCATATGCTCG 2222

QY 441 GluValCysGluAspSerIleThrValGluMetValArgHisGlnGluLysIleMetTrp 460
DB 2223 GAGGTGTGTAGAGCTTCATTAACGTGGCATGTGGTCCGGCACCGGGAAATCATGTAC 2282
QY 461 ValGluAspValArgSerValThrGlnLysHisIleGlnGluTrpGluProPheAspLeu 480
DB 2283 GTCGGGGACGTCCGACAGGTGCACAGAAAGCATTTCCAGAGTGGGGCCATTGCATCTG 2342
QY 481 ValIleGluLysTrpProCysAsnAspLeuSerIleValAsnProAlaArgLysGluLeu 500
DB 2343 GTGATTTGGGGACATCCCTGTCAATGACCTTCCATGTCCAACTGTGCTCCGAAAGGCTTC 2402
QY 501 TyrGluGluTrpArgLysLeuPhePheGluPheTrpArgLeuLeuHisAspAlaArgPro 520
DB 2403 TACAGGGACATGGCGGCTCTTCTTGAAGTTTACCGGCTTCTGTGATGATGGCGGCC 2462
QY 521 LysGluGluAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGluVal 540
DB 2463 AAGGAGGGAATGATCCGCCCTTCTTGTGGCTTTTGAAATGTGTGGCCATGGGCGTT 2522
QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
DB 2523 AGTACAAAGAGGACATCTGCGCATTTCTCGAGTCCAACTGTGATGATTGATGCCAA 2582
QY 561 GluValSerAlaAlaHisArgAlaArgTrpPheTrpGluAsnLeuProGluMetAsnArg 580
DB 2583 GAGGTGTACGTGCACAGAGGCCGCTTCTTGTGGGTAACTTCCGGGTATGAAACGG 2642
QY 581 ProLeuLaserThrValAsnAspLysLeuGluLeuGlnGluCysLeuGlnHisGluArg 600
DB 2643 CCGTTGCATTCACCTGGAATGATTAAGCTGAGGTGCAGAGGTGTCTGAGACATGGCAG 2702
QY 601 IleAlaLysPheSerLysValArgThrIleThrTrpArgSerAsnSerIleLysGlnGly 620
DB 2703 ATACCCAAAGTTCAGCAAGAGAGACCATTAACAAGGTCAACCTCCATAAACAAGGCGC 2762
QY 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
DB 2763 AAAGACCAAGATTTTCTGTCTTCTTATGAATGAGAAAGAGACATCTTATGTGTGACGAA 2822
QY 641 MetGluArgValPheGluPheProValHisTrpThrAspValSerAsnMetSerArgLeu 660
DB 2823 ATGGAAGGGATTTGGTTTCCAGTGCATTAATCTGACGCTCCACATGAGCGGCTTG 2882
QY 661 AlaArgGlnArgLeuLeuGluArgSerTrpSerValProValIleArgHisLeuPheAla 680
DB 2883 GCAGAGCAAGACTGTGCGGCGGCTCATTTGAGCGGTGCCAGTCAATCCGCACCTTCTGCT 2942
QY 681 ProLeuLysGluTrpPheAlaCysVal 689
DB 2943 CCGCTGAAGAGATTTTGTGCTGTGTG 2969

RESULT 6
US-09-720-086-3
Sequence 3, Application US/09720086
Publication No. US20060084053A1
GENERAL INFORMATION:
APPLICANT: Li, En
APPLICANT: Okano, Masaki
APPLICANT: Xie, Shaoqing
TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
TITLE OF INVENTION: Polypeptides & Uses Thereof
FILE REFERENCE: 0609,4560002
CURRENT APPLICATION NUMBER: US/09/720,086
CURRENT FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US99/14373
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/090,906
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/093,993
PRIOR FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 82
SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 3
; LENGTH: 4293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-720-086-3

Alignment Scores:
Pred. No.: 0
Score: 3763.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 5

Length: 4293
Matches: 689
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-623-813-86 (1-689) x US-09-720-086-3 (1-4293)

Qy      1  MetAenAlaValIGluGluGlnGlnGlyProGlyGluSerGlnLysValGluGluLysSer 20
Db      899  ATGAATGCTGTGGAGAAACCAAGGGGCCCGGGAGTCTCAAGAGTGGAGAGGCCAAGC 958

Qy      21  ProProAlaValGlnGlnProThraSProAlaSerProThraValAlaThraProGlu 40
Db      959  CCTCTGCTGTGACAGACCCCACTGACCCCGCATCCCACTGCTGCTACCAAGCCTGAG 1018

Qy      41  ProValGlySerAspAlaGlyAspLysAsnAlaThraLysAlaGlyAspAspGluProGlu 60
Db      1019  CCCGTGGGGTCCGATGCTGGGGACAAGATGCCCAAGACGAGCGATGACGAGCCAGAG 1078

Qy      61  TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValATrPGlyLysLeuArgGlyPhe 80
Db      1079  TACAGAGACGCGCGGGCTTTGGCATTTGGGAGCTGTGTGGGGAAACCTGGGGGCTTC 1138

Qy      81  SerTrpTrpProGlyArgGlyLeuIleSerTrpTrpMetThrGlyArgSerArgAlaIleGlu 100
Db      1139  TCCTGTGGCCGACGCGCATTTGTCTTGTGTGATGACGCGCCGAGCGCAAGCACTGAA 1198

Qy      101  GlyTrpArgTrpValMetTrpPheGlyLysAspGlyLysPheSerValValCysValGluLys 120
Db      1199  GGCACCCCTGGGTATGTGTTCCGAGCGGCAAAATCTCAGTGTGTGTGAAG 1258

Qy      121  LeuMetProLeuSerSerPheCysSerAlaPheIleGlnAlaThrTyrAsnLysGlnPro 140
Db      1259  CTGATGCGCGTGAAGCTCGTTTGGAGTGTCCACAGGCCACCTACACAAGAGGCC 1318

Qy      141  MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      1319  ATGTACCCCAAGCATCTACAGAGGTCTCGCAGGTGGCCAGCAGCGCGGGGAAGCTG 1378

Qy      161  PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValGlnAsn 180
Db      1379  TTCCTGGTGTCCAGCAGCAGCATGATGATGACACTGCGCAAGGCCGTGAGTGCAGAAC 1438

Qy      181  LysProMetIleGluTrpAlaLeuGlyIlePheGlnProSerGlyProLysGlyLeuGlu 200
Db      1439  AAGCCCAAGATTGAATGAGGCCCTGGGGGCTTCCAGCCTTGTGGCTTAAGGGCTTGAG 1498

Qy      201  ProProGluGluGluLysAsnProTyrLysGluValTyrThraSPMetTrpValGluPro 220
Db      1499  CCACCAGAAAGAGAAATCCCTCAAGAAAGTGTACAGCGAGCATGTGGTGTGAACCT 1558

Qy      221  GluAlaAlaAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db      1559  GAGGAGGCTGCTTACGACCACTCCACAGCCAAAGAGCCCGGAAAGACCAAGCGAGG 1618

Qy      241  LysProLysValLysGluIleIleAspGluArgThrArgLysArgLeuValTyrGluVal 260
Db      1619  AAGCCCAAGGTCAGAGGATTAATGATGAGGCCACAAGAGCGGCTGTGTGATGAGAGTG 1678

Qy      261  ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1679  CGGAGAAAGTCCCGGAAATTTGAGAGCATCTGCACTCTGTGGAGAGCCTCAATTTTACC 1738

Qy      281  LeuGluHisProLeuPheValGlyIleMetCysGlnAsnCysLysAsnCysPheLeuGlu 300

```

```

Db      1739  CTGAAACACCCCTTCTTGTGGAGAGATGTCCAAACTCAAGAACTCTTCTTGAG 1798

Qy      301  CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      1799  TGTGTGTACCAAGTACGAGAGAGAGCGGCTACAGGTCTTACTCAACCATCTGTGTGGGGC 1858

Qy      321  ArgGluValLeuMetCysGlyLysAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1859  GGTAGGGCTCAAGTGTGGGAAACAAACACTGTGAGGTGCTTTTGTGGAGGTGTGTG 1918

Qy      341  AspLeuLeuValGlyProGlyAlaAlaGlnAlaIleLysGluAspProTyrAsnCys 360
Db      1919  GACCTCTTGGTGGGGCCCGGGGCTTCCAGGACCACTTAAGAAAGACCTTGGAACTGC 1978

Qy      361  TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGlyAspTrpProSer 380
Db      1979  TACATGTGCGGGCACAAGAGGTACCTACGGGCTGTGGCGCGGAGAGGACTGGCCCTCC 2038

Qy      381  ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProLysValTyr 400
Db      2039  CGGCTCCAGATGTTCTTGTGTAATTAACACGACAGAAATTTGACCTTCCAAAGGTTTAC 2098

Qy      401  ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2099  CCACCTGTCCAGCTGAGAAAGAGAACCCATCCGGGTGCTGTCTCTTGTGATGGAATC 2158

Qy      421  AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db      2159  GCTACAGGGCTCTGTGTGTGAAGACTTGGGACTTCAAGGTGAGCCGCTTACATTGGCTCG 2218

Qy      441  GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2219  GAGGTGTGTAGGACTCATTCACGATGCGTGTGGCCACGAGGGAATATGATGATAC 2278

Qy      461  ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTrpGlyProPheAspLeu 480
Db      2279  CTCGGGAGCGTCCGACGTCACACAGACATTCACAGAGTGGGCCCATTTGCATCTG 2338

Qy      481  ValIleGlyLysProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2339  GTGATTTGGGGCAGTCTCTGCAATGACCTCTCCATCTCTCAACCTGTCCGACAGGCTTC 2398

Qy      501  TyrGluGlyThrGlyArgGluPhePheGluPheTyrArgLeuHisAspAlaArgPro 520
Db      2399  TACAGGGCACTGGCCGGCTTTTGTGAAGTCTACCGGCTCTGCAATGATGCCGGCCC 2458

Qy      521  LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaMetGlyVal 540
Db      2459  AAGGAGGAGATGATCCGCCCTTCTTGTGGCTCTTTGAGATGTGTGGCCATGGGCTT 2518

Qy      541  SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      2519  AGTGACAAAGAGGACATCTCGCATTTCTGAGTCCAACTGTGATGATGATGAGCCAAA 2578

Qy      561  GluValSerIleAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      2579  GAAATGTCACTGACACAGAGGCCCGCTTACTTCTGGGGTAACTTCCCGGATGAACAGG 2638

Qy      581  ProLeuAlaSerThrValAsnAspLysLeuGluGlnGluCysLeuGluHisGlyArg 600
Db      2639  CCGTTGGCATTCACCTGTGAATGATAGCTGAGCTGAGAGAGTGTCTGAGACATGGCAGG 2698

Qy      601  IleAlaLysPheSerLysValArgThrIleThrThraSerAsnSerIleLysGlnGly 620
Db      2699  ATAGCCAAAGTTACGAAAGTGAAGACCATTACTACGAGTCAACTCCATTAAGCAGAGGC 2758

Qy      621  LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
Db      2759  AAGACCAAGATTTCTGTCTTCAATGATGAAGAAAGGACATCTTAGGTGACACTGAA 2818

Qy      641  MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660

```

Db 2819 ATGAAAGGATTTGGTTCCTCCAGTCCACTATACTAGCTCTCCAAACATGAGCCGCTTG 2878
Qy 661 AATAGGTATATGLeuLeuGlyArgSerThrSerValProValIleArgHisLeuPheIa 680
Db 2879 GCGAGGACAGACGTCTGGGCGGCTCATGAGCGCTCCAGTCATCCGCCACTTCTCGCT 2938
Qy 681 ProLeuLeuGlyIleThrPheAlaCysValI 689
Db 2939 CCGCTGAAGAGATATTTTGGCTGTGTG 2965

RESULT 7
US-10-623-813-3
; Sequence 3, Application US/10623813
; Publication No. US2004023497A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; APPLICANT: Chen, Taiding
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0609, 4560003
; CURRENT APPLICATION NUMBER: US/10/623, 813
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 09/720,066
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/US99/14373
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,906
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 60/093,993
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 4293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-813-3

Alignment Scores:
Pred. No.: 0 Length: 4293
Score: 3763.00 Matches: 689
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-623-813-86 (1-689) x US-10-623-813-3 (1-4293)

Qy 1 MetCAmAlaValGluGluAsnGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
Db 899 ATGAATGCTGTGAAAGAAACAGAGGCGCCGAGAGTCTCAGAAAGTGAGAGAGCCAGC 958
Qy 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db 959 CCTCTGCTGTGACGACGCCACTGACCCGCCACTGCCACTGTGGCTACCAACCCCTGAG 1018
Qy 41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
Db 1019 CCGGTGGGCTCCGATGCTGGGGACAGAAATGCCAACAGAGCGATGACGAGCCAGAG 1078
Qy 61 TyrGluAspGlyArgGlyPheGlyIleGlyCyluLeuValTyrGlyLysLeuArgGlyPhe 80
Db 1079 TAGCAGAGACCGCCGCGGCTTTGGCATTTGGGAGCTGGTGTGGGAGAACTGGCGGCTTC 1138
Qy 81 SerTyrTrpProGlyArgGlyLeuAlaSerTyrTrpMetThrGlyArgSerArgAlaGlu 100
Db 1139 TCCGTGGTGGCAGGCCCATTTGTCTTGGTGGATGACGGGCCGAGACCAAGCTGGA 1198
Qy 101 GlyThrArgTyrValMetTyrPheGlyAspGlyLysPheSerValIleCysValGluLys 120
Db 1199 GGACCCGCTGGGTCAITGTGTGGAGACGGCAAAATTTCTCAGTGTGTGTGGAGAG 1258

Qy 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db 1259 CTGATGCCGCTGACCTGTGTTTGCAGTGCCTTCCACAGGCCACGTAACACAGAGCC 1318
Qy 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db 1319 ATGTACCGCAAGCATCTAGAGGTCTCTGAGGTGGCCAGACCGCGGGGGAAGCTG 1378
Qy 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValGluValAsn 180
Db 1379 TTCCCGGTGTGCCAGCAGAGATGAGTGCACCTCCAAAGGCCGTGGAGGTGCAGAAC 1438
Qy 181 LysProMetIleGluTyrPalaLeuGlyIlePheGlnProSerGlyProLysGlyLeuGlu 200
Db 1439 AAGCCATGATGTGATGGAGCCCTGGGGGCTTCCAGCTTCTGAGCTTAAAGGCTGGAG 1498
Qy 201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTyrValGluPro 220
Db 1499 CCACCAAGAGAGAGAGATCCCTTCAAGAGATGTACACGACATGTGGTGGAACT 1558
Qy 221 GluAlaAlaAlaTyrAlaProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db 1559 GAGGCACTGCTCAAGCACCACCTCCACGACCAAAAGCCCCGGAAGAGACAGCGGAG 1618
Qy 241 LysProLysValLysGluIleIleAspGluArgThrArgGluArgLeuValTyrGluVal 260
Db 1619 AAGCCCAAGGTCAAAGGATTTGATGAGCGCACAAAGAGCCGCTGTGTACAGAGTG 1678
Qy 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db 1679 CCGAGAGAGTCCGGGAACATTGAGAGATCTGCATCTCGTGGAGACCTCAATGTTACC 1728
Qy 281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
Db 1739 CTGGAAACACCCCTCTGTTGGAGATGTGCCAAATGCAAGAACTGCTTCTCGAG 1798
Qy 301 CysAlaIleTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGly 320
Db 1799 TGTGCTTACCAAGTACGACGACGAGGCTTACCACTCTACCTGACCACTGCTGTGGGGC 1858
Qy 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1859 CGTAGAGTGTCTCATGTGCGGAAACAACACTGCTGAGGTGCTTTGCGTGGAGTGTG 1918
Qy 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaAlaIleLysGluAspProTyrAsnCys 360
Db 1919 GACCTCTTGTGGGGCCGGGGGCTGCCACAGCACTTAAGGAAGACCCCTGGAACTGC 1978
Qy 361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgGlyAspTyrProSer 380
Db 1979 TACATGTGGGACACAGAGGTACTTACGAGCTGTGGGGGCGAGAGACTGGCCCTCC 2018
Qy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProLysValTyr 400
Db 2039 CCGGTCCAGATGTTCTTCGCTAATAACCAAGCAAGCAATTGACCTCCAAAGCTTTAC 2098
Qy 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 2099 CCACTGTCCACGTGAGAGAGAGAGCCCATCCGGGTGTCTCTCTTGAAGGAAATC 2158
Qy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 2159 GCTACAGGGCTCTCTGGGTGAAAGACTTGGCATTTAGGTGACCCCTACATTGCTCG 2218
Qy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db 2219 GAGGTGTGAGGATCTCATCAGGTGGCATGTGGGACCAAGGAGATCATGTAC 2278
Qy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnIleTyrGlyProPheAspLeu 480
Db 2279 GTGCGGAGCTCCGACGCTCACAAGAACATATCCAGAGAGTGGGGCCCATTCGATCTG 2338

```

Qy 481 ValIleGIyGlySerProCyAsnAspLeuSerIleValaenProAlaArgGlyLeu 500
Db 2339 GTGATGGGGAGTCCCTGCATGACCTCTCCATGCTCAACCTGCTGGCAAGGGCTTC 2398
Qy 501 TyrGIuGIyThrGIyArgLeuPhePheGluPheTyraGLeuLeuHisAspAlaArgPro 520
Db 2399 TACGAGGGCACTGGCGGCTCTCTTTGAGTTCTACCCGCTCTTCAGATGCGCGGCC 2458
Qy 521 LyseGIuGIyAspAspArgProPhePheTrpLeuPheGluAsnValaIleMetGIyVal 540
Db 2459 AAGGGGAGATGATGATCGCCCTCTTCTTGCGCTCTTGAAGATGCTGGCCATGGCCCTT 2518
Qy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 2519 AGTGAACAAGAGGACATCTCGCGATTTCTCGAGTCAACCTCTGTATATGATGCCAA 2578
Qy 561 GluValSerAlaAlaHisArgAlaArgTyrrPheTrpGIyAsnLeuProGIyMeCAsnArg 580
Db 2579 GAAGTGTAGCTGCACACAGGGCCCGCTACTTCTGGGTAACCTTCCCGTATGAAACAG 2638
Qy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGluGluCysLeuGluHisGIyArg 600
Db 2639 CCGTTGGCATCCACTGTGTAATGATAGCTGAGCTGCAGAGTGTCTGAGACATGGCAGG 2698
Qy 601 IleAlaLysPheSerLysValArgThrIleThrTrpArgSerAsnSerIleLeuGluGly 620
Db 2699 ATACCCAAAGTTCAGAAAGTGAAGGACCACTACTACAGAGTCAAACTCCATTAAGCAGG 2758
Qy 621 LysAspGIHisPheProValPheMetAsnGIuGIyAspIleLeuTrpCysThrGIu 640
Db 2759 AAAGACCAAGCATTTTCCGTCTTCATGAAAGAAAGAGACATCTTAAGTGCACTGAA 2818
Qy 641 MetGIuArgValPheGIyPheProValHisTyrrThrAspValSerAsnMetSerArgLeu 660
Db 2819 ATGGAAGGGATATTGGTTTCCAGTCCACTATATGACGTCTCCAAATGAGCGGCTTG 2878
Qy 661 AlaArgGIAspLeuLeuGluArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2879 GCGAGGCGAAGACTGCTGGCGCGGTCAATGAGCGGATCATTCGCCACTCTTCCCT 2938
Qy 681 ProLeuGIyGIyTyrrPheAlaCysVal 689
Db 2939 CCGCTGAAGAGTATTGCGGTGTG 2965

RESULT 8
US-10-623-813-83
; Sequence 83, Application US/10623813
; Publication No. US20040234997A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoping
; APPLICANT: Chen, Taiping
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0609.4560003
; CURRENT APPLICATION NUMBER: US/10/623.813
; PRIOR FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 09/720.086
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/US99/14373
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090.906
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 60/093.993
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83
; LENGTH: 2318
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-623-813-83

```

```

Alignment Scores:
Pred. No.: 0 Length: 2318
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatch: 8
Query Match: 98.7% Indels: 0
DB: 9 Gaps: 0

US-10-623-813-86 (1-689) x US-10-623-813-83 (1-2318)
Qy 1 MetAsnAlaValGIuGIuAsnGIuGlyProGIyGluSerGIuValGIuAlaSer 20
Db 148 ATGATGCTGTGGAAGAAGAACAGAGCCCTCGAAGATCTCGAAGGTGAGAGGAGCCAGC 207
Qy 21 ProProAlaValGIuGluInProThrAspProAlaSerProThrValAlaThrThrProGIu 40
Db 208 CTTCTGCTGTGCGAGCGCCACCGACCTGTCTCTCGACTGTGGCCACCCCTGAG 267
Qy 41 ProValGIySerAspAlaGIyAspLysAsnAlaThrLysAlaGIyAspAspGIuProGIu 60
Db 268 CCACTAGAGAGGGATGCTGGGACAAAGATGCTTCCAAAGCAGCCGAGATGAGCTTGA 327
Qy 61 TyrGIuAspGIyArgGIyPheGIyIleGIyGluLeuValITrpGIyLysLeuArgGIyPhe 80
Db 328 TATGAGATGGCCGGGCTTTGGCATTGAGAGCTGTGTGGGAAACTTCGGGGCTTC 387
Qy 81 SerTrpTrpProGIyArgIleValSerTrpTrpMetThrGIyArgSerArgAlaAlaGIu 100
Db 388 TCCGTGTGGCCAGGCCCAATTTGTCTTGTGTGATGACAGGCCGAGCCGAGAGCTGAA 447
Qy 101 GlyThrArgTrpValMetTrpPheGIyAspGIyLysPheSerValaIleCysValGIyLys 120
Db 448 GGCACGTGCTGGGTCACTGTGTTGCGAGATGCAAGTTCAGTGTGTGTGTGAGAG 507
Qy 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGIuAlaThrTyraSnLysGIuPro 140
Db 508 CTCATGGCGGTGAGCTCTCTCTGAGTGCATTCCACAGGCGCACTTAACAAACAGAGCC 567
Qy 141 MetTyrrArgLysAlaIleTyrrGIuValLeuGluValaIleSerSerArgAlaGIyLysLeu 160
Db 568 ATGTACCGCAAGCATCTAGAGTCTCCAGGTGCGAGCAGCCGTCGGGAAAGCTG 627
Qy 161 PheProValCysHisAspSerAspGIuSerAspThrAlaLysAlaValaIleGIuAlaSn 180
Db 628 TTTTCACTTGCCATGTGACATGATGAAGTGAACGTGCAAGCTGTGGAAGTCAAGAC 687
Qy 181 LysProMetIleGIuTrpAlaLeuGIyGIyPheGluProSerGIyProLysGIyLeuGIu 200
Db 688 AAGCAGATGATGAATGGGCCCTCGGTGGCTTCCAGCCCTCGGTCTTAAGGGCTGAG 747
Qy 201 ProProGIuGIuGIyLysAsnProTyrrLysGIuValTyrrThrAspMetTrpValGIuPro 220
Db 748 CCAACCAAGAAAGAGAAATCTTCAAGAAATTTAACCCGACAGTGGTGGAGCT 807
Qy 221 GluAlaIleAlaTyrrAlaProProProProAlaLysLysProArgLysSerThrAlaGIu 240
Db 808 GAAGCAGCTGCTTACGCCCAACCCCAAGCCAGAAAGCCAGAAAGAGCACAACAGAG 867
Qy 241 LysProLysValLysGIuIleIleAspGIuArgThrArgLysLeuValTyrrGIuVal 260
Db 868 AAACCTAAGGTCAAGAGATCATTTGATGAGGCAAGAGGCGGCTGTGTATGAGGTG 927
Qy 261 ArgGIuLysCysArgAsnIleGIuAspIleCysIleSerCysGIySerLeuAsnValThr 280
Db 928 CCGCAGAAAGTGCAGAAACATCGAGGACATTGTATCTCATGTGGAGCCCTCAAGTCAAC 987
Qy 281 LeuGIuHisProLeuPheValaIleGIyLysCysGIuAsnCysLysAsnCysPheLeuGIu 300
Db 988 CTGAGACACCACTCTTCAATTGAGGATGTGCCAGAACTGTAAAGACTGTTTGGAG 1047
Qy 301 CysAlaTyrrGIuTyrrAspAspAspGIyTyrrGlnSerTyrrCysThrIleCysCysGIyGIy 320

```

Db 1048 TGTGCTTACCAAGTATGACGACGATGGGTACCAAGTCTTATGACCAATCTGCTGGGGG 1107
Qy 321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db 1108 CGTAAAGTCTCAATGCTGGGAAACAACATGCTGCAAGTCTTTGTGTGAGTGTG 1167
Qy 341 AspleuLeuValGlyProGlyAlaAlaGlnAlaAlaIleValGluAspProTyrAsnCys 360
Db 1168 GATCTCTTGGTGGGCGACGAGCTGCTCAGCAGCAGCATTAGAAAGACCCCTGGAACTGC 1227
Qy 361 TyrMetCysGlyHisIleValGlyThrTyrGlyLeuLeuArgArgGluAspTyrProSer 380
Db 1228 TACATGTCCGGGCAATAAGGACCACTATGGCTGCTGCACAAACGGGAAGACTGGCTCTT 1287
Qy 381 ArgLeuLeuMetPhePheAlaAsnAsnHisAspGlnGluPheAspProTyrValTyr 400
Db 1288 CGACTCCAGATGTTCTTTGCAATACCAATGACCAAGAAATTGACCCCCCAAGGTTTAC 1347
Qy 401 ProProValProAlaGluValArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db 1348 CCACTGTGCCAGCTGAGAAAGAGAACCCCATCCCGCTGCTCTCTTGTGAGGAT 1407
Qy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAlaSer 440
Db 1408 GCTACAGGGCTCCCTGCTGTAAGACCTGGGCACTCCAAAGTGAACCGCTACATTGCCCTCC 1467
Qy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyValIleMetTyr 460
Db 1468 GAGGTGTGAGAGCTCATCAACGCTGAGGACGCTGGGCAACGAGGAAGATCATGTAC 1527
Qy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrGlyLysProPheAspLeu 480
Db 1528 GTGGGGAGCTCCCGACGCTCAACAGACATATCCAGAGTGGGGCCATTGACCTG 1587
Qy 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db 1588 GTGATTTGAGAGGAGTCCCTCAATGACCTGCTCCATGTGCAACCTGCCCAAGGACTT 1647
Qy 501 TyrGluGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
Db 1648 TATGAGGGTACTGACCCCTCTTTGTGAGTCTACCGCTCCGCTGATGTCGGGGCCC 1707
Qy 521 LysGluGlyAspAspArgProPhePheTyrPhePheGluAsnValValAlaMetGlyVal 540
Db 1708 AAGGAGGAGATGATGCCCTCTTCTGCTCTTGAAGATGCTGGCGCATGGGCTT 1767
Qy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db 1768 AGTGACAAAGAGGACATCTCGCATTTCTTGAGCTTAACCCCGTGATGATTGACGCCAA 1827
Qy 561 GluValSerAlaAlaHisArgAlaArgTyrPheTyrGlyAsnLeuProGlyMetAsnArg 580
Db 1828 GAAGTGTGCTCTCAACAGGGCCGCTTACTTCTGGGGTAACTTCTGGCATGAACAGG 1887
Qy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGlyCysLeuGluHisIleArg 600
Db 1888 CCTTGGCATCCACTGTAATGATTAAGCTGAAGCTGCAAGCTCTGGAGCAGCGAGA 1947
Qy 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db 1948 ATAGCCAAAGTTCAAGAAAGTGAAGACCATTAACCAACGAGTCAAACTGTATAAGCAGG 2007
Qy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
Db 2008 AAAGACCAAGATTTCCCGCTTCAATGAACGAAGAGGAGCAATCTGTGTGACAGAA 2067
Qy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2068 ATGGAAAGGGTGTGTTGGCTTCCCGCTCACTACAGACGCTTCCAAACAGACGGCTTG 2127
Qy 661 AlaArgGlnArgLeuGlyArgSerTyrSerValProValIleArgHisLeuPheAla 680
Db 2128 GCGAGGCAAGACTGCTGGGCGCATGTTGAGGCGTGCCTCATCCGCCACTTCTGCT 2187

Qy 681 ProLeuLysGluTyrPheAlaCysVal 689
Db 2188 CCGCTGAAGAAATTTTCTGCTGTGTG 2214

RESULT 9
US-09-720-086-1
; Sequence 1, Application US/09720086
; Publication No. US20060084053A1
; GENERAL INFORMATION:
; APPLICANT: Li, En
; APPLICANT: Okano, Masaki
; APPLICANT: Xie, Shaoqing
; TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
; FILE REFERENCE: 0609,456002
; CURRENT APPLICATION NUMBER: US/09/720,086
; PRIOR FILING DATE: 2000-12-20
; PRIOR FILING DATE: 1999-06-25
; PRIOR FILING DATE: 1999-06-25
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090,906
; PRIOR APPLICATION NUMBER: 60/093,993
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4192
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (4161)..(4161)
; OTHER INFORMATION: May be any nucleic acid
; US-09-720-086-1

Alignment Scores:
Pred. No.: 0 Length: 4192
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.7% Indels: 0
Gaps: 0
US-10-623-813-86 (1-689) x US-09-720-086-1 (1-4192)

Qy 1 MetAsnAlaValGluGluLeuGlnGlyProGlyGluSerGlnLysValGluGluAlaSer 20
Db 874 ATGAATGCTGTGGAAGAAACAGGCTCTGGAGAGTCTCAGAAAGTGGAGAGCCAGC 933
Qy 21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrTyrProGlu 40
Db 934 CTTCTGCTGTGAGAGAGCCAGGACCTGCTTCCGCTGTGCTGACACACCCCTGAG 993
Qy 41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
Db 994 CCAAGTGAAGGAGATCTGGGACCAAGATCTCAACCAAGCAGCCAGATGAGCTGAG 1053
Qy 61 TyrGluAspGlyValArgGlyPheGlyIleGlyGluLeuValTyrGlyLysLeuArgGlyPhe 80
Db 1054 TATGAGAGATGGCGCGGCTTTGGCATTTGAGAGCTGTGTGGGGAAACTTCCGGGCTTC 1113
Qy 81 SerTyrTyrProGlyValArgIleValSerTyrTyrMetThrGlyValSerArgAlaIleGlu 100
Db 1114 TCTGTGTGCGCAGGCGGAATGTGTCTTGTGTGATACAGGCGGAGCGAGCGCTGAA 1173
Qy 101 GlyThrArgTyrValMetTyrPheGlyAspGlyLysPheSerValValCysValGluLys 120
Db 1174 GCGACCTGGTGGGTACTGTGTGTGCGGATGCGCAATTTCTCACTGTGTGTGTGGAAG 1233
Qy 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db 1234 CTATATCCCTGAGAGTCTTCTTGTGCAAGTGCATTCACAGGCGCACTTAACAAAGCAGCCC 1293

Oy 141 MetYrAglYsAlaIleTyrGlValLeuGlnValAlaSerSerAglAglYsLeu 160
 Db ATGTACCCCAAGCATCTACCAAGCTCTCCAGGTGGCCAGCAGCCGTGCCGGGAAGCTG 1353
 Oy 161 PheProValCySHisAspSerAspGluSerAspThrAlaValGluValGlnAsn 180
 Db TTTCAGCTTGCCAGCAGTGTATGAAGTACAGCAGGTGTGAGAGTGCAGAAC 1413
 Oy 181 LysProMetIleGluTrpAlaLeuGlyPheGlnProSerGlyProGlyLeuGlu 200
 Db AAGCAGATGATTGAATGGGCCCTCGGTGGCTTCCAGCCCTGGGTCTTAAGGGCTTGAG 1473
 Oy 201 ProProGluGluGluValAsnProTyrIleGlyValIleTyrAspMetTyrValGluPro 220
 Db CCACCAAGAAAGAAAGAAATCTTACAGAGAGTTTACCCGACATGTGGTGGAGCT 1533
 Oy 221 GluAlaAlaIleTyrAlaProProProAlaValLysLysProArgLysSerThrAlaGlu 240
 Db GAAGCAGCTGCTTACGCCCCCACCACCGCAAGAAACCAAGAAAGCACAAGAG 1593
 Oy 241 LysProLysValLysGluIleIleAspGluArgThrArgLysValTyrGluVal 260
 Db AAACCTAAGCTCAAGAGATCATTCATGAGCGCACAAGAGCGCGCTGTGTAGAGCTG 1653
 Oy 261 ArgGlnLysCyAspArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
 Db CGCCAGAAAGTCAGAAACATCGAGACATTTGTATCTCATGTGGAGCCTCAATGTCCAC 1713
 Oy 281 LeuGluHisProLeuPheValGlyGlyMetCysGlnAsnCysLysAsnCyPheLeuGlu 300
 Db CTGGAGCACCACCTTTCATTTGAGGAGCATGTGCCAACTGTAAACCTGCTTCTTGAG 1773
 Oy 301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGly 320
 Db TGTGCTTACCAATGACGAGCATGTGGTACCAAGTCTATGACCATCTGCTGTGGGG 1833
 Oy 321 ArgGluValLeuMetCysGlyLysAsnAsnCysCysArgCyPheCysValGluCysVal 340
 Db CGTGAAGTCTCATGTGTGGGAAACAACATGCTGCAGAGTCTTTGTGTGAGGTGTG 1893
 Oy 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaIleLysGluAspProTrpAsnCys 360
 Db GATCTCTTGTGGGCCAGAGCTCTCTAGGAGCCATTAAAGAAAGCCCTGGAAGCTGC 1953
 Oy 361 TyrMetCysGlyHisLysGlyThrTyrIleLeuLeuArgArgGluAspTrpProSer 380
 Db TACATGTGCGGCGATTAAGGACCTATGGCTGTGGAAAGCGGGAAGACTGGCTTCT 2013
 Oy 381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnGluPheAspProProLysValTyr 400
 Db CGACTCCAGATGTTCTTTGCAATTAACATGACCGAGAAATTGACCCCAAGGGTTTAC 2073
 Oy 401 ProProValProAlaGluLysArgLysProIleArgValLysSerLeuPheAspGlyIle 420
 Db CCACCTGTGCCAGCTGAGAAAGAAAGCCCATCCCGTGTCTCTCTTTTGAATGGAGATT 2133
 Oy 421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAspArgTyrIleAsnSer 440
 Db GCTACAGGGCTCTGTGTCTGAAGACCTGGGCACTCAAGTGAACCGCTACATTGCCCTCC 2153
 Oy 441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
 Db GAGGTGTGTGAGACTCATACCGTGGCATGTGTGGCCACCGAGAAATCATCTGTAC 2253
 Oy 461 ValGlyAspValArgSerValThrGlnLysHisIleGlnGluTyrPGLProPheAspLeu 480
 Db CTCGGGAGACGTCCGAGGCTCACACAGAACATATCCAGAGTGGGCGCATTTGCACTG 2313
 Oy 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
 Db GTGATTGAGAGCAGCTCTGCAATGACCTTCATCTTCAACCTCGCCGCAAGGAGACTT 2373

Oy 501 TyrGluGlyThrGlyArgLeuPheGluPheTyrArgLeuLeuHisAspAlaArgPro 520
 Db TATAGGGTACTGGCCCTCTTCTTTAGTTCTACCGCTCTGCAATGATGCCGGGCC 2433
 Oy 521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
 Db AAGAGGGAGATGATGCCCTTCTTCTTGTGCTCTTTAGATGATGTGGCCATGGGCTT 2493
 Oy 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
 Db AGTACCAAGAGGGACATCTCGCATTTCTTGAGTCTTAACCCCGATGATGATGAAGCAAA 2553
 Oy 561 GluValSerAlaIleHisArgAlaArgTyrPheTrpGlyLysLeuLeuProGlyMetAsnArg 580
 Db GAAAGTCTGCTGACACAGAGGCCGTTACTTCTGGGGTAACTTCTTGGCATAAACAGG 2613
 Oy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnGlyCysLeuGluHisGlyArg 600
 Db CTTTGGCATCTCACTGGAATGATTAAGCTGAGCTGCAGAGTGTCTGAGCAGCGCAGA 2673
 Oy 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
 Db ATAAGCAAGTTCAAGCAAGTGAAGACCATTAACAACAGTCAACTCTATTAAGCAGAGGC 2733
 Oy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTyrCysThrGlu 640
 Db AAAGACCAAGCTTCCCGCTTCTATGAAGAGAGAGAGACATCTGTGTGATGACACTGA 2793
 Oy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
 Db ATGGAAGGGGTGTTGGCTTCCCGCTCACTACACAGACGCTCAACATGACGCGCTTG 2853
 Oy 661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
 Db GCGAGCGAGAGACTGCGGGCCGATCGTGAAGCGTGCCTGATCCGCCACTTCTTGCT 2913
 Oy 681 ProLeuLysGluTyrPheAlaCysVal 689
 Db CCGCTGAAGAAATTTTGTCTGTGTG 2940

Db 2914 CCGCTGAAGAAATTTTGTCTGTGTG 2940

RESULT 10
 US-10-623-813-1
 : Sequence 1, Application US/10623813
 : Publication No. US20040234997A1
 : GENERAL INFORMATION:
 : APPLICANT: Li, En
 : APPLICANT: Okano, Masaki
 : APPLICANT: Xie, Shaoping
 : APPLICANT: Chen, Taiping
 : TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes, Polypeptides & Uses
 : FILE REFERENCE: 0609, 4560003
 : CURRENT APPLICATION NUMBER: US/10/623, 813
 : CURRENT FILING DATE: 2003-07-22
 : PRIOR APPLICATION NUMBER: US 09/720, 086
 : PRIOR FILING DATE: 1999-06-25
 : PRIOR APPLICATION NUMBER: PCT/US99/14373
 : PRIOR FILING DATE: 1999-06-25
 : PRIOR APPLICATION NUMBER: US 60/090, 906
 : PRIOR FILING DATE: 1998-06-25
 : PRIOR APPLICATION NUMBER: US 60/093, 993
 : PRIOR FILING DATE: 1998-07-24
 : NUMBER OF SEQ ID NOS: 119
 : SOFTWARE: PatentIn version 3.2
 : SEQ ID NO 1
 : LENGTH: 4192
 : TYPE: DNA
 : ORGANISM: Mus musculus
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (4161)..(4161)
 : OTHER INFORMATION: n is a, c, g, or t
 US-10-623-813-1

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Pred. No.: | 0 | Length: | 4192 |
| Score: | 3714.00 | Matches: | 679 |
| Percent Similarity: | 98.9% | Conservative: | 2 |
| Best Local Similarity: | 98.5% | Mismatches: | 8 |
| Query Match: | 98.7% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-10-623-813-86 (1-689) x US-10-623-813-1 (1-4192)

```
QY      1 MetAsnAlaValGluGluAsnGlnGlyProGlyGlySerGlnLysValGluGluAlaSer 20
Db      874 ATGAATGCTGTGGAAGAAACCAAGCCCTGTGAAGTCTCAAGAAAGTGAAGAGCCAGC 933

QY      21 ProProAlaValGlnGlnProThrAspProAlaSerProThrValAlaThrThrProGlu 40
Db      934 CTTCTCTGCTGGCAGCAGCCACCGAACCTCTCTCCGACTGTGGCCACCACTCCCTGAG 993

QY      41 ProValGlySerAspAlaGlyAspLysAsnAlaThrLysAlaGlyAspAspGluProGlu 60
Db      994 CCAATAGAGAGGGAGTGTGGGGACAAAGATCTCAACAAAGCAGCCGAGATGAGCTGAG 1053

QY      61 TyrGluAspGlyArgGlyPheGlyIleGlyGluLeuValTrpGlyLysLeuArgGlyPhe 80
Db      1054 TATGAGATGGCCGGGCTTTGGCATTGAGAGCTGTGTGGGGAAACTTCGGGGCTTC 1113

QY      81 SerTrpTrpProGlyArgGlyLeuAlaSerTrpMetThrGlyArgSerArgAlaAlaGlu 100
Db      1114 TCTCGTGGCCAGGCCCAATGTGTCTGTGTGATGACAGGCCGGAAGCCGAGAGCTGAA 1173

QY      101 GlyThrArgTrpValMetTrpPheGlyAspGlyLysPheSerValValCysValGluLys 120
Db      1174 GGCACTGGCTGGGTCATGTGTGTGAGATGGCAAGTCTCAGTGTGTGTGTGAGAG 1233

QY      121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
Db      1234 CTGATGGCGGTGAGCTCTTCTGCAGTGCATTCCACAGGCCACTTCACAAACAGAGCC 1293

QY      141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
Db      1294 ATGTACCGCAAGCATCTCAAGAGTCTCCAGGTGGCCAGCAGCCCTGCGGGAGAGCTG 1353

QY      161 PheProValLysHisAspSerAspGlySerAspThrAlaLysAlaValGluValGluAsn 180
Db      1354 TTTTCAGCTTGCATGACGATGAAGTAAGTAGACAGTGGCAAGCTGTGGAGTGCAGAAC 1413

QY      181 LysProMetIleGluTrpAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
Db      1414 AAGCAGATGATTGAATGGCCCTCGGTGGCTTCCAGCCCTCGGGTCCCTAAGGGCTGGAG 1473

QY      201 ProProGluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrpValGluPro 220
Db      1474 CCACCAAGAAAGAGAAATCTTACCAAGAGATTTCACCGACATGTGGTGGAGAGCT 1533

QY      221 GluAlaIleAlaTyrAlaProProProProAlaLysLysProArgLysSerThrAlaGlu 240
Db      1534 GAAGCAGCTCTTAAGCCCAACCCCAAGCCCAAGCAAAACCAAGAAAGAGCAAAACAGAG 1593

QY      241 LysProLysValLysGluIleIleAspGlyLysArgThrArgGluArgLeuValTyrGluVal 260
Db      1594 AAACCTAAGGTCAAGAGATCATTTGATGAGGCAAAAGGAGCGGGCTGTGTATAGAGTG 1653

QY      261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
Db      1654 CGCAGAAAGTGCAGAAACATCTGAGAGCATTTGTATCTCATGTGGAGACTCAAGTCAACC 1713

QY      281 LeuGlnHisProLeuPheValGlyGlyMetCysGlnLysCysLysAsnCysPheLeuGlu 300
Db      1714 CTGAGAGCAACCACTTCTTGGAGCATGTGCAAGAACTGTAAAGAACTGTTTGGAG 1773

QY      301 CysAlaTyrGlnTyrAspAspAspGlyTyrGlnSerTyrCysThrIleCysCysGlyGly 320
Db      |||
```

```
Db      1774 TGTCTTACCAAGTATGACGAGATGGGTACCAAGTCTATTGCAACCATCTGCTGGGGG 1833
QY      321 ArgGluValLeuMetCysGlyAsnAsnAsnCysCysArgCysPheCysValGluCysVal 340
Db      1834 CGTAATGTCTCAATGTGTGGGAACAACATCTGCAAGGTCTTTGTGTGAGTGTGTG 1893

QY      341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaIleLysGluAspProTrpAsnCys 360
Db      1894 GATCTCTTGGTGGGGCCAGAGAGCTGTCAAGGAGCCATTAAAGAAAGACCCCTGGAATGCG 1953

QY      361 TyrMetCysGlyHisLysGlyThrTyrGlyLeuLeuArgArgArgGluAspTrpProSer 380
Db      1954 TACATGTGGCGGCATTAAGGACACTTATGGCTGTGTGGAAGACCGGAAGACTGGCTCT 2013

QY      381 ArgLeuGlnMetPhePheAlaAsnAsnHisAspGlnLysPheAspProLysValTyr 400
Db      2014 CGACTCCAGATGTTCTTTGCAATTAACATAGACCAAGAAATTTGACCCCAAGGTTTAC 2073

QY      401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
Db      2074 CCACTGTGCCAGCTGAGAAAGAGAACCCATCCGCTGTCTCTCTTGTGATGGAGATT 2133

QY      421 AlaThrGlyLeuLeuValLeuLysAspLeuGlyIleGlnValAlaAspArgTyrIleAlaSer 440
Db      2134 GCTACAGAGCTCTCTGTGTGAAGGACCTGGGCATCCAAAGTGACCCGCTACATTGCCCTCC 2193

QY      441 GluValCysGluAspSerIleThrValGlyMetValArgHisGlnGlyLysIleMetTyr 460
Db      2194 GAGGTGTGTAGGACTTCATCAGCGTGGCATGTGTGGCACCAGGAAGAAATATATGAC 2253

QY      461 ValGlyAspValArgSerValThrGlnLysHisIleGlnLysTrpGlyProPheAspLeu 480
Db      2254 GTCGGGAGCTCCGCAAGCTCAACAGAACATTTCCAGAGTGTGGCCATTCGACCTG 2313

QY      481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
Db      2314 GTGATTGAGAGGAGTCCCTCAATGACCTCTCCATTGTCAACCTGCGCCGCAAGGGAATT 2373

QY      501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuHisAspAlaArgPro 520
Db      2374 TATGAGGTACTGGCCCTCTTTGAGTTCACCGCTCTGACATGATGAGTGGCCATGGGCTT 2433

QY      521 LysGluGlyAspAspArgProPhePheTrpLeuPheGluAsnValAlaIleMetGlyVal 540
Db      2434 AAGGAGGAGATGATGCCCTCTTCTGTGCTCTTTGAGATGTGTGTGGCCATGGGCTT 2493

QY      541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
Db      2494 AGTGACAAAGAGGACATCTCGCATTTCTTGAGCTTAACCCCGGTGATGATTGAGCCCAA 2553

QY      561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGlyAsnLeuProGlyMetAsnArg 580
Db      2554 GAAGTGTCTCTGACACAGAGGCCCTTACTTCTGGGGTAACTTCTGGGATTAACAGG 2613

QY      581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGlnLysCysLeuGluHisGlyArg 600
Db      2614 CTTTGGCATCCACTGTGAATGATAGCTGAGCTGCAAGATGTCTGAGAGCAGGAGAG 2673

QY      601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGlnGly 620
Db      2674 ATAGCCAAAGTTACAGAAAGTAGGACCATTAACCAACCGGTCAAACTCTATAAGCAGAGGC 2733

QY      621 LysAspGlnHisPheProValPheMetAspGluLysGluAspIleLeuTrpCysThrGlu 640
Db      2734 AAAGACCAAGATTTCCCGTCTTCAATGAACAGAAAGAGGACATCTGTGTGTGACAGTA 2793

QY      641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db      2794 ATGGAAGAGGTGTTGGCTTCCCGTCCACTACACAGACGCTTCCACATGAGCGGCTTG 2853

QY      661 AlaArgGlnArgLeuLeuGlyArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db      2854 GCGAGGACAGACTGTGGGCGATGTGAGAGGTGCGGTGTCACTCCGCACTCTTGCT 2913
```

QY 681 ProteulysgluTyrrhealaCysVal 689
DB 2914 CCGCTGAAGAAATATTTGCTGTGTG 2940

RESULT 11
US-10-330-773-276
; Sequence 276 Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 6157
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-330-773-276

Alignment Scores:
Pred. No.: 0 Length: 6157
Score: 3714.00 Matches: 679
Percent Similarity: 98.8% Conservative: 2
Best Local Similarity: 98.5% Mismatches: 8
Query Match: 98.7% Indels: 0
Gaps: 0

US-10-623-813-86 (1-689) x US-10-330-773-276 (1-6157)

QY 1 MetasnaIaValIGluIAsnGInglYProglYIuSerGInLysValIGluIuAlaSer 20
DB 866 ATGAATGCTGTGAAGAAACACAGGCCTCTGGAGAGTCTCAAGAGTGGAGAGCCAGC 925

QY 21 ProProAlaValIGInLInProThrsPProAlaSerProThrsAlaThThProglu 40
DB 926 CCTCTGCTGTGCAGACCCACAGGACCTCTCTCCGACTGTGGCCACACCCCTGAG 985

QY 41 ProValIGlySerAspAlaGlyAspLysAsnaIaThrLysAlaGlyAspAspGluProglu 60
DB 986 CCAAGTACAGAGGCGATGCTGGGGAACAAGATCTACCAAGACAGCCGAGAGAGCTGAG 1045

QY 61 TyrGluAspGlyArgGlyPheGlyTLeGlyGluLeuValTrrGlyLysLeuArgGlyPhe 80
DB 1046 TATGAGGATGGCCGGGCTTGGCATTTGAGAGCTGTGTGGGGAACCTTCGGGGCTTC 1105

QY 81 SerTrrPrrProglYArgGlyLeValSerTrrPrrMetThrGlyArgSerArgAlaAlaGlu 100
DB 1106 TCCGCTGTGGCCAGGCGCAATTGTCTTGTGTGAGAGACAGGCGGAGCGAGCTGAA 1165

QY 101 GlyThrArgTrrValMetTrrPheGlyLysPheSerValValCysValIGluLys 120
DB 1166 GGCACCTGCTGGGTCATGTGTTCCGAGATGGCAAGTTCTCAGTGTGTGTGTGAGAG 1225

QY 121 LeuMetProLeuSerSerPheCysSerAlaPheHisGlnAlaThrTyrAsnLysGlnPro 140
DB 1226 CTCATGCCGCTGAGCTCTCTCTGAGTGCATTCCACAGGCCACCTCAACAAGAGGCC 1285

QY 141 MetTyrArgLysAlaIleTyrGluValLeuGlnValAlaSerSerArgAlaGlyLysLeu 160
DB 1286 ATGTAACGCAAGCCATCTACGAAGTCTCCAGAGTGGCCAGAGCGCTGGGGAAGCTG 1345

QY 161 PheProValCysHisAspSerAspGluSerAspThrAlaLysAlaValIGluAsn 180
DB 1346 TTTTCAAGCTTCCATGACAGTGAATAAGTGAACAGTGCAGAGCGTGGAGTGCAGAAC 1405

QY 181 LysProMetIleGluTrrAlaLeuGlyGlyPheGlnProSerGlyProLysGlyLeuGlu 200
DB 1406 AAGCAGATGATTGAATGGGCTTCGTTCCAGCTTCGGGCTCTTAAGGGCTTGAG 1465

QY 201 ProProgluGluGluLysAsnProTyrLysGluValTyrThrAspMetTrrValIGluPro 220
DB 1466 CCACAGAAAGAGAAATCTCTTACAGAGAACTTACCCAGCATGTGGTGGAGCTT 1525

QY 221 GluAlaAlaIaTrrAlaProProProProAlaLysPheArgLysSerThrAlaGlu 240
DB 1526 GAACACAGCTGTTACGCCGCCACCCACAGCCAAAGAAACCCAGAAAGAGACAAACAG 1585

QY 241 LysProLysValLysGluIleIleAspGluArgThrArgGluValTyrGluVal 260
DB 1586 AAACCTAAGGTCAGAGGATCATTTGATGAGCGCAAGAGGAGCGCTGTGTAGAGGTG 1645

QY 261 ArgGlnLysCysArgAsnIleGluAspIleCysIleSerCysGlySerLeuAsnValThr 280
DB 1446 GCCAGAGTGCAGAAACATCGAGGACATTTGATCTCATGTGGAGACCTCAATGTAC 1705

QY 281 LeuGluHisProLeuPheValIGlyMetCysGlnAsnCysLysAsnCysPheLeuGlu 300
DB 1706 CTGGAGCACCACTCTCATTTGAGGAGCATGTGCCAGAACTGTAAAGAACTGCTTCTGAG 1765

QY 301 CysAlaTyrGlnTrrAspAspAspGlyTrrGlnSerTrrCysThrIleCysCysGlyGly 320
DB 1766 TGTGCTTACAGATGACGAGAGGAGTACCATCTTATTCACCATCTGCTGTGGGGG 1825

QY 321 ArgGluValLeuMetCysGlyAsnaAsnaCysCysArgCysPheCysValIGluCysVal 340
DB 1826 CGTCAAGTCTCATGTGTGGGAACAAACATCTGCAGAGTGTCTTGTGTGATGTGTG 1885

QY 341 AspLeuLeuValGlyProGlyAlaAlaGlnAlaIleLysGluAspProTrrAsnCys 360
DB 1886 GATCTCTTGGTGGGCGCAGAGAGCTGCTCAGGACCACTTAAGAAAGCCCTGTGAACCTGC 1945

QY 361 TyrMetCysGlyHisLysGlyThrTrrGlyLeuLeuArgArgArgLysAspTrrProSer 380
DB 1946 TACATGTGCGGCAATGAAGGACCTATAGGCTGTGTGAAGACCGGAAGACTGGCTTCT 2005

QY 381 ArgLeuGlnMetPhePheAlaAsnaAsnHisAspGlnGluPheAspProLysValTyr 400
DB 2006 CGACTCAGATGTTCTTGTGCATTAACATGACAGGAATTGACCCCCAAAGGTTTAC 2065

QY 401 ProProValProAlaGluLysArgLysProIleArgValLeuSerLeuPheAspGlyIle 420
DB 2066 CCACCTGTCCAGCTGAGAGAGAGAGCCCATCGCTGTCTCTCTTGAAGGAGATT 2125

QY 421 AlaThrGlyLeuLeuValIleLysAspLeuGlyIleGlnValAspArgTrrIleAlaSer 440
DB 2126 GCTACAGAGGCTCTCTGTGTGAAGGACTGGGCACTCCAAAGTGAACCTTACATTGCCCTCC 2185

QY 441 GluValCysGluAspSerIleThrValIGlyMetValArgHisGlnGlyLysIleMetTyr 460
DB 2186 GAGGTGTGAAGACTCATCACGCTGGGCACTGTGGGCAACAGGAAAGATATATAC 2245

QY 461 ValGlyAspValArgSerValIleGlnLysHisIleGlnGluTrrGlyProPheAspLeu 480
DB 2246 GTGGGAGACGCCGACAGCTCACAGAAACATATCCAGAGTGGGCGCCATTGACCTG 2305

QY 481 ValIleGlyLysSerProCysAsnAspLeuSerIleValAsnProAlaArgLysGlyLeu 500
DB 2306 GTGATTTGAAGAGCTCTCTCAATGACCTTCACTGTCAACCTTGCCCGAAGGAGCTT 2365

QY 501 TyrGluGlyThrGlyArgLeuPhePheGluPheTyrArgLeuHisAspAlaArgPro 520
DB 2366 TATACAGGTACTGGCCGCTCTTCTTGAATTTCTACCGGCTCTCGCATGATGGCGGCC 2425

QY 521 LysGluGlyAspAspArgProPhePheTrrPhePheGluAsnValAlaIleMetGlyVal 540
DB 2426 AAGGAGGAGATGATCGCCCTTCTTGTGCTCTTGAAGATGTGGGCGCATGGGCTT 2485

QY 541 SerAspLysArgAspIleSerArgPheLeuGluSerAsnProValMetIleAspAlaLys 560
DB 2486 AGTACAAAGAGGACATCTCGCAATTTCTTGAATCTAACCCCGTATGATTGACGCCAAA 2545

Qy 561 GluValSerAlaAlaHisArgAlaArgTyrPheTrpGluAsnLeuProGluMetAsnArg 580
Db 2546 GAATGTCCTGCTGACACAGGGCCCGTTACTTCTGGGTAACCTTCCGGCATGAACAG 2605
Qy 581 ProLeuAlaSerThrValAsnAspLysLeuGluLeuGluGlnCysLeuGluHisGlyArg 600
Db 2606 CCTTGGCATCTCACTGAAATGATAGCTGGAGCTGCAGAGTGTCTGGACAGCGGAGA 2665
Qy 601 IleAlaLysPheSerLysValArgThrIleThrThrArgSerAsnSerIleLysGluGly 620
Db 2666 ATAGCCAAAGTTCCACCAAGTAGAGCACTTACCAACAGGTCAAACTATATAACAGAGGC 2725
Qy 621 LysAspGlnHisPheProValPheMetAsnGluLysGluAspIleLeuTrpCysThrGlu 640
Db 2726 AAGAGCCAGCATTTCCCGCTTTCATGACAGAAAGAGGACATCTGTGTGACATGAA 2785
Qy 641 MetGluArgValPheGlyPheProValHisTyrThrAspValSerAsnMetSerArgLeu 660
Db 2786 ATGGAAAGGGGTGTGGCTTCCCGCTCCACTACACAGACGTCTCCAAACATGAGCCGCTTG 2845
Qy 661 AlaArgGlnArgLeuGluArgSerTrpSerValProValIleArgHisLeuPheAla 680
Db 2846 GCGAGCGAGAGACTGCTGGGCCGATCGTGAGCGTGCGGTCAATCCGCACCTTGTGCT 2905
Qy 681 ProLeuLysGluTyrPheAlaCysVal 689
Db 2906 CCGCTGAAGGAATATTTTCTTGTGTG 2932

RESULT 12
US-10-264-237-953
; Sequence 953, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Btse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; PCT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 953
; LENGTH: 2938
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (491)..(491)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1987)..(1987)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1994)..(1994)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2892)..(2892)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2932)..(2932)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-953

Alignment Scores: 2,11e-293 Length: 2938
Pred. No.: 2713.00 Matches: 496

Percent Similarity: 99.0% Conservative: 0
Best Local Similarity: 99.0% Mismatches: 5
Query Match: 72.1% Indels: 0
DB: 7 Gaps: 0

US-10-623-813-86 (1-689) x US-10-264-237-953 (1-2938)

Qy 189 GlyGlyPheGlnProSerGlyProLysGlyLeuGluProProGluGluLysAsnPro 208
Db 1 GGCACGAGCCAGCCTTCTGTGCTTGAAGGCTTGAAGCCACACAGAAAGAAATCCC 60
Qy 209 TyrLysGluValTyrThrAspMetTrpValGluProGluAlaAlaTyrAlaProPro 228
Db 61 TACAAAGATGTAACAGGACATGTGGGTGAACCTGAGGACAGCTGCTACGACACACT 120
Qy 229 ProProAlaLysLysProAlaGlySerThrAlaGluLysProLysValLysGluIleIle 248
Db 121 CCACAGCCCAAAAAGCCCGGAAAGACACGCGAAGAGCCCAAGTCAAGGAGATTATT 180
Qy 249 AspGluArgThrArgGluArgLeuValTyrGluValArgLysCysArgAsnIleGlu 268
Db 181 GATGAGGCAACAGAGAGCGGCTGTGTACGAGGTGCGGCAAGTCCGGAACATTGAG 240
Qy 269 AspIleCysIleSerCysGlySerLeuAsnValThrLeuGluHisProLeuPheValGly 288
Db 241 GACATCGCATCTCTGTGGAGCTTCAATGTTACCTGGAAACCCCTCTTGTGGA 300
Qy 289 GlyMetCysGlnAsnGlyLysAsnGlyPheLeuGluCysValTyrGlnTyrAspAsp 308
Db 301 GGAATGCGCAAAACGTGCAGAACTGTTTGTGAGTGTGACCTACGACGACGACGAC 360
Qy 309 GlyTyrGlnSerTyrCysThrIleCysCysGlyGlyArgGluValLeuMetCysGlyAsn 328
Db 361 GCTTACAGTCTTACTGACCATCTGCTGTGGGGCCGTAGAGTCTCACTGCGGAAC 420
Qy 329 AsnAsnGlyCysArgCysPheCysValGluCysValAspLeuValGlyProGlyAla 348
Db 421 AACACTGCTGACGGTCTTTCGTTGAGTGTGACCTCTTGGTGGGGCGGGGCT 480
Qy 349 AlaGlnAlaAlaIleLysGluAspProTrpAsnGlyMetCysGlyHisLysGlyThr 368
Db 481 GSCCARCAGNCATTAAAGAAAGCCCTGGAACTGCTACATGTGGGGCAAGAGGTAAC 540
Qy 369 TyrGlyLeuLeuArgArgArgGluAspTrpProSerArgLeuGlnMetPhePheAlaAsn 388
Db 541 TACGGGTGCTGGGGCGGCGAGAGACTGGCCTCCGGTCCAGATCTTCCGTAAT 600
Qy 389 AsnHisAspGlnLysPheAspProProLysValTyrProProValProAlaGluLysArg 408
Db 601 AACCAAGACCAAGAAATTGACCTCCAAAGTTTACCACTGCTCCAGCTGAAGAGG 660
Qy 409 LysProIleArgValLeuSerLeuPheAspGlyIleAlaThrGlyLeuLeuValLeuLys 428
Db 661 AAGCCATCCGGGTGCTCTCTTTGATGGAATGCTACAGGCTCTGTGTGTAAG 720
Qy 429 AspLeuGlyIleGlnValAspArgTyrIleAlaSerGluValCysGluAspSerIleThr 448
Db 721 GACTTGGGCAATTAGGTGACCCCTACATTGCTCGAAGTGTGTAGACTCCATCAGC 780
Qy 449 ValGlyMetValArgHisGlnGlyLysIleMetTyrValGlyAspValArgSerValThr 468
Db 781 GTGGGATGTGCGGCACCGAGGGAAGATCATGTACCTCGGGAACGTCACAGCTGACA 840
Qy 469 GlnLysHisIleGlnGluTyrProPheAspLeuValIleGlyGlySerProCysAsn 488
Db 841 CAGAAGCATATCCAGAGAGGGGCCCATTCGATGTGTATGGGGCAGTCCCTGCAAT 900
Qy 489 AspLeuSerIleValAsnProAlaArgLysGlyLeuTyrGluGlyThrGlyArgLeuPhe 508
Db 901 GACTCTCCATCTCAACCTGCTCGAAGGCTCTACAGAGGCACTGGCCGGCTTTC 960
Qy 509 PheGluPheTyrArgLeuLeuHisAspAlaArgProLysGluGluLysAspAspArgPhe 528

```

Db      961 TTGAGCTTACCGCCTCTGCATGATGCGCGCCCAAGAGGAGATGATGCCCTTC 1020
Oy      529 PheTTPLeuPheGluAsnValValAlaMetGlyValSerAspLysArgAspLLeuSerArg 548
Db      1021 TTCTGGCTCTTTGAAATAGTGGTGGCCATGGGCGTTAGTACAAAGAGGAGACATCTCCGGA 1080
Oy      549 PheLeuGluSerAsnProValMetLeuAspAlaLysGluValSerAlaAlaHisArgAla 568
Db      1081 TTTCTCGAGTCCAACTCTGTGATGATGATGTCCAAGAAAGTGTCACTGCACACAGGGCC 1140
Oy      569 ArgTyrPheTTPGlyAsnLeuProGlyMetAsnArgProLeuAlaSerThrValAsnAsp 588
Db      1141 CGCTACTCTGGGGTCAACTCTCCGTTATGAACAGCGCGTTGGCATCTTCATGTGATGAT 1200
Oy      589 LysLeuGluLeuGlnGluCysLeuGluHisGlyValGluLeuAlaLysPheSerLysValArg 608
Db      1201 AAGCTGGAGCTGCAGAGAGTGTCTGGAGCATGGCAGAGATAGCCAACTTACAGCAAGTGAAG 1260
Oy      609 ThrLeuThrThrArgSerAsnSerLLeuLysGlnGlyLysAspGlnHisPheProValPhe 628
Db      1261 ACCATTACTACGAGGTCAAACCTCCATPAAGCAGGCGCAAGACACAGCAATTTTCTCTTC 1320
Oy      629 MetAsnGluLysGluAspLLeuLeuTTPCysThrGluMetGluArgValPheGlyPhePro 648
Db      1321 ATGAATGAGAAAGAGCATCTTATGTCGCACTGAATGAAAGGCTATTGGTTTCCCA 1380
Oy      649 ValHisTyrThrAspValSerAsnMetSerArgLeuAlaArgGlnArgLeuLeuGlyArg 668
Db      1381 GTCCACTTACTGACGTCTCCCAACATGAGCCGCTGGCGAGGCAAGACTGTGGGCGCG 1440
Oy      669 SerTTPSerValProValLeuArgHisLeuPheAlaProLeuLysGluTyrPheAlaCys 688
Db      1441 TCATGAGAGCGTCCAGTCAATCCGCACTCTTCGCTCGCTGAAGAGATATTTCGCTGT 1500
Oy      689 Val 689
Db      1501 GTG 1503

```

```

RESULT 13
US-09-720-086-2
Sequence 2, Application US/09720086
Publication No. US20060084053A1
GENERAL INFORMATION:
APPLICANT: Li, En
APPLICANT: Okano, Masaki
APPLICANT: Xie, Shaoping
TITLE OF INVENTION: De Novo DNA Cytosine Methyltransferase Genes,
FILE REFERENCE: 0609 4560002
CURRENT APPLICATION NUMBER: US/09/720,086
CURRENT FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US99/14373
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/090,906
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/093,993
PRIOR FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 82
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 4195
TYPE: DNA
ORGANISM: Mus musculus
US-09-720-086-2

```

```

Alignment Scores:
Pred. No.: 4,92e-232 Length: 4195
Score: 2169.50 Matches: 404
Percent Similarity: 71.94 Conservative: 100
Best Local Similarity: 57.68 Mismatches: 156
Query Match: 57.74 Indels: 41
DB: 5 Gaps: 6

```

```

US-10-623-813-86 (1-689) x US-09-720-086-2 (1-4195)
Oy      10 ProGluGluSerGlnLysValGluGluAlaSerProProAlaValGlnGlnProThrAsp 29
Db      797 CTTCGACGCGTCGACTTCAAGAGAAAGTAGACACTTAAGAGCGGTCACTACCCCA----- 850
Oy      30 ProAlaSerProThrValAlaThrThrProGluProValGlySerAspAlaGlyAspLys 49
Db      851 -----TCAGTTGACTTGAGCCAGCAGATGAGATGAGAGGGTATGATATACCAACAGGTG 904
Oy      50 AsnAlaThrLysAlaGlyAspAspGluProGluTyrGluAspGlyArgGlyPheGlyTyr 69
Db      905 GATCGACAGAGACAGATGAGATGAGACAGACAGATATCAGATATGAAGAGTTTGGATATA 964
Oy      70 GlyGluLeuValTTPGlyLysLeuArgGlyPheSerThrTTPProGlyArgGlyLeuValSer 89
Db      965 GGTGACCTCGTGGGGAAGATCAAGGGCTTCTCCGTGGTGGCTGCCATGTGTGTTC 1024
Oy      90 TTPThrMetThrGlyArgSerArgAlaAlaGluGlyThrArgTTPValMetTTPPheGly 109
Db      1025 TGGAAACCCACCTCCACAGCAGCAGCCATGCCGGAATGCCGTGGTATCAGTGTGGT 1084
Oy      110 AspGlyLysPheSerValValCysValGluLysLeuMetProLeuSerSerPheCysSer 129
Db      1085 GATGCAAGTTTCTGTGAGATCTGTGACAAACTGTGGTGTGGGCTGTTCAGCCAG 1144
Oy      130 AlaHisGlnAlaThrTyrAsnLysGlnProMetTyrArgLysAlaLeuTyrGluVal 149
Db      1145 CACTTATCTGGTATCTTCAATAACTGGTGTCTTATAGAAAGGCCATGTACACACT 1204
Oy      1205 CTGAGAAAGCCAGGCTTCAGAGCTGCGAAGACCTC-----TCC 1243
Db      170 SerAspThrAlaLysAlaValGluValGlnAsnLysProMetLLeuTTPAlaLeuGly 189
Oy      1244 AGCAGTCTCGAGAGTCACTGAGAGAACCACTGAAGCCATGTGGTGGGCCACAGGT 1303
Oy      190 GlyPheGlnProSerGlyProLysGlyLeuGluPro----- 201
Db      1304 GGCTTCAAGCTTACTGGATGAGGGGCTCAACCCAAAGAAAGCAACAGTGTGAAT 1353
Oy      202 -----ProGluGluGlnLysAsnPro 208
Db      1364 AAGTCGAAGGTGCGTCTGACAGATGAGAACTTAGACCCAGGAGACGAGAACAAA 1423
Oy      209 TyrLysGluValTyrThrAspMetTTPValGluProGluAlaAlaTyrAlaProPro 228
Db      1424 AGTCGAAGACGCAACAACCAAT-----GACTCTGCTCTTGTGACTCCCC 1468
Oy      229 ProProAlaLysLysProArgLysSerThrAlaGluLysProLysValLysGluLLeu 248
Db      1469 CCACCCCAAGGCGCTCAAGACAAATAGCTATAGCGGGAAGACCGAGGGAG----- 1519
Oy      249 AspGluArgThrArgGluArgLeuValTyrGluValArgGlnLysCysArgAsnLLeuGlu 268
Db      1520 GATAGAGAGACCGAAGACCGATGCGCTTCAAGTCAACAACAAGGGAATCTCGAA 1579
Oy      269 AspLLeuCysLLeuSerCysGlySerLeuAsnValThrLeuGluHisPProLeuPheValGly 288
Db      1580 GACCGCTGTTGCTCTGTGAAAGAAAGAACCTGTGCTTCCACCCCTCTTTGAGGGT 1639
Oy      289 GlyMetCysGlnGlnGlnCysLysAsnCysPheLeuGluCysAlaTyrGlnTyrAspAsp 308
Db      1640 GGGCTCTGTCAAGTTGCGGGAGTGGCTTCTCAAGGCTCTTCAATGATGATGAGAGAC 1699
Oy      309 GlyTyrGlnSerTyrCysThrLLeuCysCysGlyGlyArgGluValLeuMetCysGlyAsn 328
Db      1700 GGCTATCAGTCTACTCACTCAACGCTGTGTGAGGGCCGTGAACCTGCTGTGAGATAC 1759
Oy      329 AsnAsnCysCysArgCysPheCysValGluCysValAspLeuLeuValGlyProGlyVala 348
Db      1760 ACAAGCTGTGCAAGATGCTTCTGTGTGAGATGTCTGAGAGTGTGTGGGCGCAGGACA 1819

```



```

US-10-623-813-86 (1-689) x US-09-720-086-4 (1-4145)
QY 13 SerGlnLysValIGluGlnLaseProProAlaValIGlnProThraspProAlaser 32
   |||
   |||
   |||
Db 632 ACACAGAGACACACACTGGGACGCCCAAGACGACGTACCCCTTACG---CCGCGCTTAC 688
QY 33 ProThraValAlaThr-----ThrProGluProValGlySerAspAlaGlyAsp 48
   |||
   |||
   |||
Db 689 CCCAGAGACACGACGAGGGGGGATGGAGTCCCGGAG-GTGGAGGACACAGTGGAGAT 747
QY 49 LysAsnAlaThrLysAlaGlyAspAspGluProGluTyGluAspGlyArgGlyPheGly 68
   |||
   |||
   |||
Db 748 -----GGAGACAGTTCAGACTATCAGAGTGGAGAGGATTGGAA 786
QY 69 IlegLysLysValITroGlyLysLeuAspGlyPheSerITroProGluArgLysLeuAl 88
   |||
   |||
   |||
Db 787 ATAGGGACCTGTGTGGGAAAGATCAAGGCTTCTCGGTGGGCGCCCACTGGGTGGT 846
QY 89 SerITroProMetThrGlyArgSerArgAlaAlaGluGlyIThrArgITroValMetITroPhe 108
   |||
   |||
   |||
Db 847 TCTTGGAAAGCCCACTCCAGAGGACAGGCTATGTCTGGAGTGGGTGGTCCAGTGGTTT 906
QY 109 GlyAspGlyLysPheSerValValGlyValGlyLysLeuMetProLeuSerSerPheCys 128
   |||
   |||
   |||
Db 907 GGCAGATGGCAAGTCTCCAGAGTCTCTCGACAAACTGGTGGCACTGGGCGCTTTCAGC 966
QY 129 SerAlaPheHisGlnAlaIThrTyraAsnLysGlnProMetITyrArgLysAlaITyrGlu 148
   |||
   |||
   |||
Db 967 CAGCACTTAAATTGGCCACTTCATTAAGCTGCTCTCTATCGAAAGCCATGTACAT 1026
QY 149 ValLeuGlnValAlaSerSerArgAlaGlyLysLeuPheProValCysHisAspSerAsp 168
   |||
   |||
   |||
Db 1027 GCTCTGGAGAAAGCTTAGGGTGGAGCTGGCAAGACTTCC----- 1068
QY 169 GluSerAspIThrAlaLysAlaValGluValGlnAsnLysProMetIlegLysITroAlaLeu 188
   |||
   |||
   |||
Db 1069 ---AGCACCCCTGGAGACTCATTTGGAGGACACAGCTGAAGCCCATTTGGAGTGGGCCCCAC 1125
QY 189 GlyGlyPheGlnProSerGlyProLysGlyLeuGluProProGluGlu----- 204
   |||
   |||
   |||
Db 1126 GGGGGCTTCAAGCCACTGGGATCGAGGGCTTCAACCCACACACAGCCAGCAGTGGTT 1185
QY 205 -----GluLysAsnProITyrLysGlu 211
   |||
   |||
   |||
Db 1186 AATAAGTCGAAGCTCCGTCGACAGGAGTAGAATTAAGATCAAGAAATACGAAAC 1245
QY 212 ValITyrIThrAspMetITroValGluProGluAlaAlaAla---TyrAlaProProPro 230
   |||
   |||
   |||
Db 1246 AAGACTCGAAGACGACAGCTGACGACTCAGCCACCTCTGACTACTGCCCCGACCCCAAG 1305
QY 231 AlaLysLysProITroArgLysSerIThrAlaGluLysProLysValLysGluLeuLeuAspGlu 250
   |||
   |||
   |||
Db 1306 CGCCTCAAGACAAATTGCTATTAAACAAGCGAAAGACGAGGGAGTAA-----GAT 1356
QY 251 ArgIThrArgGluArgLeuValITyrGluValAlaArgGlnLysCysArgAsnIlegLysAspIle 270
   |||
   |||
   |||
Db 1357 CAGACCCAGAAACAATGGCTTACAGATGTGGCCACACACAAAGACAGCTGGAAATGGC 1416
QY 271 CysIleSerCysGlySerLeuAsnValIThrLeuGlnHisProLeuPheValGlyGlyMet 290
   |||
   |||
   |||
Db 1417 TGTGTGCTCTGTGGCAGGAAAAAACCCTGCTCTTCCACCCCTCTTGGAGGGGGGGCT 1476
QY 291 CysGlnAsnCysLysAsnCysPheLeuGluCysAlaITyrGlnITyrAspAspAspLysITyr 310
   |||
   |||
   |||
Db 1477 TGTACAGACATCCGGGATCGCTTCTTGAAGCTGTTTACATGATTAAGATACATGGCTAT 1536
QY 311 GlnSerITyrCysIThrIleCysCysGlyArgGluValAlaLeuMetCysGlyLysAsnAsn 330
   |||
   |||
   |||
Db 1537 CAGCTTACTGCACTGTGTGTCTGGAGGGCCGAGAGCTGTCTTGGACAGCAACAGAC 1596
QY 331 CysCysArgCysPheCysValGluCysValAlaPheLeuLeuValGlyProGlyAlaAlaGln 350
   |||
   |||
   |||
Db 1597 TGTGTCCCGTGTCTGTGTGTGGATGCTCGAGAGTGTCTGTGGACAGGCGACAGCGGCC 1656

```

```

QY 351 AlaAlaIleLysGluAspProITroAsnCysITyrMetCysGlyLysLeuGlyIThrTyrgly 370
   |||
   |||
   |||
Db 1657 GAGGCCAAGCTTCAGAGACCCCTGGAGCTGTCTACATGTGTCTCCGACAGGCTGTATGGC 1716
QY 371 LeuLeuArgArgArgGluAspITroProSerArgLeuGlnMetPhePheAlaAsnAsnHis 390
   |||
   |||
   |||
Db 1717 GTCTGCGGCGCCGAAAGAGATGGAACGTGGCCCTCGACAGGCTTCTTCCACAGTGCACAG 1776
QY 391 AspGlnGluPheAspProProLysValITyrProProValProAlaGluLysAspITro 410
   |||
   |||
   |||
Db 1777 GGGCTTGAATAGAAAGCCCAAGCTGTACCTGCAATCCGCGACGCCGAAAGGGGGCC 1836
QY 411 IleArgValIleSerLeuPheAspGlyIleAlaIThrGlyLeuLeuValLeuLysAspLeu 430
   |||
   |||
   |||
Db 1837 ATTCAGAGCTCTCTACTTGTGTGATGGCATCGGACAGGCTACTGTCTTCAAAAGATTG 1896
QY 431 GlyIleGlnValAspArgITyrIleAlaSerGluValCysGluAspSerIleIThrValGly 450
   |||
   |||
   |||
Db 1897 GGCATTAAGATGGAAGATGACGTCGCTTCTGAAGTGTGAAGAGATCCATGTCTGTGGA 1956
QY 451 MetValArgHisGlnGlyLysIleMetITyrValGlyAspValArgSerValIThrGlnLys 470
   |||
   |||
   |||
Db 1957 ACCGTGAAGCACAGAGGGAATATCAATACGTGAACGACGTGAAGACATTCACAAAGAA 2016
QY 471 HisIleGlnGluITroGlyProPheAspLeuValIlegLysITroCysAsnAspLeu 490
   |||
   |||
   |||
Db 2017 AATATTGAAGATGGGCGCCATTGTGATTTGGATGGCGGAAGCCCATGACATTC 2076
QY 491 SerIleValAsnProAlaArgLysGlyLeuITyrGluGlyIThrGlyArgLeuPheGlu 510
   |||
   |||
   |||
Db 2077 TCAAAATGTGAATCCAGCCAGGAAAGCCTGTATGAGGTTACAGGCGGCTTCTTCCAA 2136
QY 511 PheITyrArgLeuLeuHisAspAlaArgProLysGluGlyAspAspArgProPheITro 530
   |||
   |||
   |||
Db 2137 TTTTACCACTCGCTGAATTACTCAGCCGCCAAGAGGGTGTATGACGGCGCTTCTTCCG 2196
QY 531 LeuPheGluAsnValAlaAlaMetGlyLysAspLysArgAspIleSerArgPheLeu 550
   |||
   |||
   |||
Db 2197 ATGTTTGAAGATTTGTGTAGCCATGAAGTTGGCGCAAGAGGACATCTCACGTTCCG 2256
QY 551 GluSerAsnProValMetIleAspAlaLysGluValSerAlaAlaHisArgAlaArgITyr 570
   |||
   |||
   |||
Db 2257 GAGTGTAAATCCAGATGATGATGATGCCATCAAGTTTCTGTGCTCACAGGCGCCGATAC 2316
QY 571 PheITroGlyAsnLeuProGlyMetAsnArgProLeuAlaSerIThrValAspAspLysLeu 590
   |||
   |||
   |||
Db 2317 TTCTGGGGCAACTAACCGGAGATGAACAGGCCCGTGAATGACATCAAGAAATGATTAATC 2376
QY 591 GluLeuGlnGluCysLeuGlnHisIlegLysITroGlyIleAlaLysPheSerITroValIThrIle 610
   |||
   |||
   |||
Db 2377 GAGCTGCAAGACTGCTTGAATACATAGATAGCAAGTTTAAAGAAATACAGACAAAT 2436
QY 611 ThrIThrArgSerAsnSerIleLysGlnLysAspGlnHisPheProValPheMetAsn 630
   |||
   |||
   |||
Db 2437 ACCACCAAGTGAATCGATCAACAGAGGGAATAACAATTTTCCCTGTGTCTGTAAT 2496
QY 631 GluLysGluAspIleLeuITroCysIThrGluMetGluArgValPheGlyPheProValHis 650
   |||
   |||
   |||
Db 2497 GGCAAAGAAAGATGTTTGTGTGTGACTGAGCTCGAAAGATCTTGGCTTCTCTGTGAC 2556
QY 651 TyrIThrAspValIleSerAsnMetSerArgLeuAlaArgGlnITroValLysITroPhe 670
   |||
   |||
   |||
Db 2557 TACACAGAGCTGTCCACATGGGCGGTGTGTGCGCCCGCAAGACTGTCTGGAAAGTCTCTGG 2616
QY 671 SerValProValIleArgHisLeuPheAlaProLeuLysGlyITyrPheAlaCys 688
   |||
   |||
   |||
Db 2617 AGCGGCTGTCAATCCAGACCTTTCGCGCCCTGGAAGACTACTTGTGATGT 2670

```

Search completed: October 6, 2006, 04:08:05
Job time : 2168 secs